MANIFOLD-CONSTRAINED EMBEDDINGS FOR THE DETECTION OF WHITE MATTER LESIONS IN BRAIN MRI

Samuel Kadoury[⋆], *Guray Erus*[†], *Evangelia I. Zacharaki*[‡], *Nikos Paragios*[∗], *Christos Davatzikos*[†]

[⋆]Ecole Polytechnique de Montreal, Montreal, Canada
[†]SBIA, Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA
[‡]School of Medicine, University of Patras, Greece
[∗]Center for Visual Computing, Ecole Centrale Paris, France

ABSTRACT

Brain abnormalities such as white matter lesions (WMLs) are not only linked to cerebrovascular disease, but also with normal aging, diabetes and other conditions increasing the risk for cerebrovascular pathologies. Obtaining quantitative measures which assesses the degree or probability of WML in patients is important for evaluating disease burden, and for evaluating its progression and response to interventions. In this paper, we introduce a novel approach for detecting the presence of WMLs in periventricular areas of the brain using manifold-constrained embeddings. The proposed method uses locally linear embedding (LLE) to create "normality" distributions in 12 locations of the brain where deviations from the manifolds are estimated by calculating geodesic distances along locally linear planes in the embedding. A smooth mapping function approximating the relationship between ambient and manifold spaces as a joint distribution maps unseen test images in the intrinsic space. We create a set of low-dimensional embeddings from 876 patches of healthy tissue in 73 subjects and test it on 396 patches imaging both WML and healthy areas in 33 subjects with diabetes. Experiments highlight the need of nonlinear techniques to learn the studied data with detection rates over 85% in true-positives, and the relevance of the computed distance for comparing individuals to a specific pathological pattern.

***Index Terms*—** Manifold embedding, locally linear embedding, white matter lesions, brain MRI, diabetes.

1. INTRODUCTION

The prevalence of cerebral white matter lesions (WMLs) has been extensively studied for the past decade in the elderly population, patients with cardiovascular risk factors or from patients with central nervous system disorders, including cerebrovascular disease and dementia. With such an increased interest towards the development of improved diagnosis and prognosis possibilities for patients with cardiovascular symptoms, specific biologic markers linking the extent of WML to the severity of the the pathology have prevailed enormous promise. In this respect, the use of an automated detection and segmentation method for WMLs exhibiting high sensitivity and specificity, which are demonstrated in a quantitative and objective fashion, could be advantageous to confidently distinguish between healthy and diseased tissue.

Towards this end, precise measurement of such pathology from magnetic resonance imaging (MRI), and more importantly measuring the evolution of a pathology over time, becomes a crucial aspect for detecting and monitoring the disease, as well as for evaluating the risk of progression. While a number of protocols are being researched to best visualize these lesions, WML are shown as hyperintense regions in fluid-attenuated inversion recovery (FLAIR) T2 images. Consequently, clarity of the lesions located in periventricular areas is increased in FLAIR images compared with other images.

In the literature, a wealth of methods have been used to segment multiple sclerosis lesions from FLAIR images using support vector machine classifier, fuzzy inference or Bayesian classification [1]. However, few have aimed to demonstrate the relevance of describing each pathological pattern as a deviation from healthy brain images along a manifold structure, allowing the computation of an appropriate distance between individuals and each pathological pattern. The definition of an optimal space for the comparison of different populations was already addressed by dimensionality reduction techniques, such as principal component analysis (PCA), Kernel-PCA or linear discriminant analysis [2]. The main limitation of the above mentioned methods lies on the assumption that the high dimensionality and complex data falls onto a linear distribution. This is not necessarily suited for the non-linear underlying structure of images of the brain cortex and where the local similar topology variations are not particulary taken into account. Manifold learning techniques intrinsically take under consideration the inherent geometry representation, and allow relevant comparison of individuals to the studied population through a low-dimensional Euclidean space map. Manifold representation has been applied for segmentation, physiological testing [3] or respiratory compensation [4]. In [5], the authors assessed morphological brain variations by calculating distances in manifold space based on metrics related to reconstruction errors, but did not treat the problem of divergence from normality.

In this work, we propose a novel framework which extends manifold learning techniques to help discriminate brain lesions from various normal tissue image profiles established as patterns of normal appearances in healthy patients (Fig. 1). These multiple non-linear embeddings obtained from a set of training images are used to create "normality" distributions where the distances of new mapped images are considered as "deviations" from healthy areas. Towards this end, we present a smooth mapping function estimating the relationship between ambient and manifold spaces as a joint distribution to map unseen test images. A collection of pairwise locally linear patches linking the new point finally estimates the geodesic distance from the normal manifold distribution. The originality of our method resides in the training on healthy tissue images, which allows to detect the presence of WML in test images, and the definition of physiological difference from normal images within manifold space. We present WML detection results on FLAIR MRI scans of elderly patients with diabetes obtained in a clinical study.

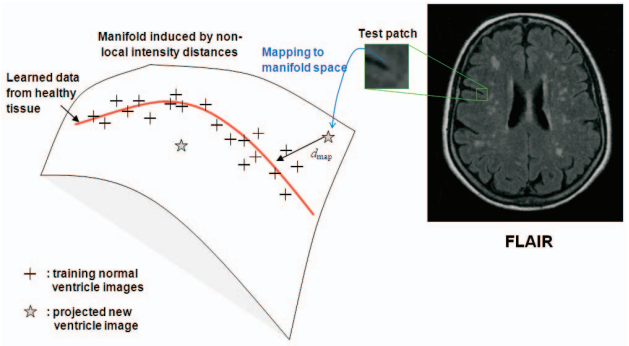


Fig. 1. Principle of the manifold-based white matter lesion detection method, embedding new sample points in the low-dimensional space to estimate their deviation from a learned healthy tissue distribution.

2. METHODS

The computation of the pathologic deviation from normality, which is measured as a distance between unseen image patches and a given population that represents the distribution of the healthy brain anatomy, consists of two steps: (1) the embedding of multiple high-dimensional datasets of ventricle images using a recently proposed weighted graph similarity measure to create locally linear manifold spaces, followed by (2) the projection of new unseen image patches onto the appropriate manifold space and determine the geodesic of the low-dimensional point to the tangential plane of the manifold.

2.1. Manifold Embedding of Periventricular Images

Let us consider a set of N images in the ambient space $\Re^D$ such that $\mathcal{I}_k = \{\mathbf{I}_1, ..., \mathbf{I}_N\} \in V_k$, extracted from a given periventricular area V_k where each image patch $\mathbf{I}_i$ is of dimensionality D. The aim is to create a low-dimensional manifold consisting of N points Y_i, $Y_i \in \Re^d$, $i \in [1, N]$ where $d \ll D$ based on locally linear embedding [6]. In such a framework, if an adequate number of data points is available, then the underlying manifold $\mathcal{M}_k$ is considered to be "well-sampled". Therefore, it can represent the underlying healthy population structure. In the sub-cluster corresponding to a pathological population, each individual point of the training set and its neighbours would lie within a locally linear patch on the manifold.

To select the K closest neighbor for each image patch, we adopt an approach similar to a fuzzy block matching approach which avoids the constraint of a strict one-to-one pairing [7]. One can assume over the image domain $\Re^D$, a weighted graph w_{ij} that links together pixels $\mathbf{x}$ of the input image $\mathbf{I}_i$ and pixels $\mathbf{y}$ of the image $\mathbf{I}_j$ with a weight $w_{ij}(\mathbf{x}, \mathbf{y})$, which is computed as follows:

$$w_{ij}(\mathbf{x}, \mathbf{y}) = \frac{1}{Z_{ij,\mathbf{x}}} e^{-\frac{\sum_{\mathbf{x}' \in P_{\mathbf{I}_i}(\mathbf{x}), \mathbf{y}' \in P_{\mathbf{I}_j}(\mathbf{y})} (\mathbf{I}_i(\mathbf{x}') - \mathbf{I}_j(\mathbf{y}'))^2}{2m\beta\sigma^2}} \quad (1)$$

where $P_{\mathbf{I}_i}(\mathbf{x})$ is a 2D ROI of the image $\mathbf{I}_i$ centered at pixel $\mathbf{x}$ and the parameter $Z_{ij,\mathbf{x}}$ is a constant of normalization such as: $\forall \mathbf{x} \in \Re^D, \sum_{\mathbf{y} \in \Re^D} w_{ij}(\mathbf{x}, \mathbf{y}) = 1$; m is the number of pixels of the image ROI; σ is the standard deviation of the noise and β is a smoothing parameter. This weighted graph w_{ij} is a representation of non-local interactions between the input image $\mathbf{I}_i$ and the image $\mathbf{I}_j$ of the training dataset. In the context of ventricular image regions in the brain, the location of the brain structures is not highly variable and it is then not desirable to go through the entire image domain to find a good match. In this work, we constrain the search of non-local structures by using a limited number of pixels around $\mathbf{x}$: $w = \{w(\mathbf{x}, \mathbf{y}), \forall \mathbf{x} \in \Re^D, \mathbf{y} \in \mathcal{S}(\mathbf{x})\}$, where $\mathcal{S}(\mathbf{x})$ is the surrounding region of the pixel $\mathbf{x}$. The size of the considered surrounding region $\mathcal{S}$ is linked to variability in the periventricular region. Hence, an ambient similarity measure is defined as: $d_{\mathbb{A}}(\mathbf{I}_i, \mathbf{I}_j) = \sum_{\forall \mathbf{x} \in \mathbf{I}_i, \mathbf{y} \in \mathcal{S}_{\mathbf{I}_j}(\mathbf{x})} \max_j w_{ij}(\mathbf{x}, \mathbf{y})$, thereby reflecting the global interactions between $\mathbf{I}_i$ and $\mathbf{I}_j$ by summing the most significant weights for each surroundings $\mathcal{S}$ in $\mathbf{I}_j$.

The manifold reconstruction weights are estimated by assuming that the local geometry of the neighborhood patches in $\mathbf{I}_i$ can be described by linear coefficients that permit the reconstruction of every model point from its neighbours. The value of the weights are therefore determined by minimizing the following reconstruction errors:

$$\varepsilon(W) = \sum_{i=1}^{N} \left\| \mathbf{I}_i - \sum_{j=1}^{N} W_{ij} \mathbf{I}_j \right\|^2 \quad (2)$$

$$\text{subject to } \begin{cases} W_{ij} = 0 \text{ if } \mathbf{I}_i \text{ not neighbor } \mathbf{I}_j \\ \sum_j W_{ij} = 1 \text{ for every } i. \end{cases} \quad (3)$$

Here, $\varepsilon(W)$ sums the squared distances between all data points and their corresponding reconstructed points. The weights W_{ij} represent the importance of the j^{th} data point to the reconstruction of the i^{th} element. It is easy to show that each weight can be calculated individually [6]. Each sample $\mathbf{I}_i$ and weight W_{ij} contributes to ε:

$$\varepsilon^{(i)}(W) = \sum_{j=1}^{K} \sum_{r=1}^{K} W_{ij} W_{ir} Q^i_{jr} \quad (4)$$

with Q as a $K \times K$ matrix, where each element is defined as $Q^i_{jr} = (\mathbf{I}_i - \mathbf{I}_j)^T (\mathbf{I}_i - \mathbf{I}_r)$ with r as an individual image region taken from the K samples of $\mathcal{I}_k$. W_{ij} are solved by a least squares problem given the constraint in (3). The algorithm maps each high-dimensional $\mathbf{I}_i$ to a low-dimensional Y_i. These internal coordinates are found with a cost function minimizing the reconstruction error:

$$\Phi(Y) = \sum_{i=1}^{N} \left\| Y_i - \sum_{j=1}^{K} W_{ij} Y_j \right\|^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} M_{ij} Y_i^T Y_j \quad (5)$$

with M as a sparse and symmetric $N \times N$ matrix enclosing the reconstruction weights W_{ij} such that $M = (I - W)^T (I - W)$, and Y spanning the Y_i's. By constraining Y to identity, the problem becomes a straightforward one with $\min \text{tr}(YMY^T)$ given the constraint $\frac{1}{N} YY^T = I$. The optimal embedding, up to a global rotation, is obtained from the bottom $d + 1$ eigenvectors of M and helps to minimize the cost function $\Phi(Y)$ as a simple eigenvalue problem.

2.2. Normality Distances of Manifold Projected Images

To obtain the low-dimensional manifold position of a new image from the ambient space (image domain), one has to infer intrinsic coordinates based on it's neighborhood representation in high-dimensional space. We first assume an explicit mapping $f : \Re^D \rightarrow \mathcal{M}_k$ from the ambient space $\Re^D$ to the manifold space $\mathcal{M}_k$. The forward mapping of $\mathbf{I}_i$ is then performed by estimating the relationship

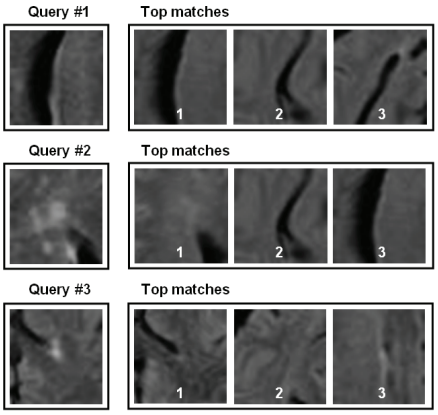


Fig. 2. Results from the manifold assignment scheme based on the weighted graph block matching method. It compares the test image to each of the 12 manifold cluster medoids.

between $\Re^D$ and $\mathcal{M}_k$ as a joint distribution, such that there exists a smooth functional which belongs to a local neighborhood. Theoretically the manifold should follow the conditional expectation:

$$f(\mathbf{I}_i) \equiv E(\mathbf{Y}|\mathcal{I}_k = \mathbf{I}_i) = \int Y \frac{p(\mathbf{I}_i, Y)}{p_{\mathcal{I}_k}(\mathbf{I}_i)} dY \tag{6}$$

which captures the overall trend of the data in d-space. Here, both $p_{\mathcal{I}_k}(\mathbf{I}_i)$ (marginal density of $\mathcal{I}_k$) and $p(\mathbf{I}_i, Y)$ (joint density) are unknown. Using *Nadaraya-Watson* kernel regression, densities are replaced by kernel functions under a conditional expectation setting [8]. We replace the variable with $p_{\mathcal{I}_k}(\mathbf{I}_i) = \frac{1}{K}\sum_{j\in\mathcal{N}(i)} G_h(\mathbf{I}_i, \mathbf{I}_j)$ and $p(\mathbf{I}_i, Y) = \frac{1}{K}\sum_{j\in\mathcal{N}(i)} G_h(\mathbf{I}_i, \mathbf{I}_j) G_g(Y, Y_j)$. The Gaussian regression kernels G require the neighbors Y_j of $j \in \mathcal{N}(i)$ to determine the bandwidths h, g so it includes all K data points ($\mathcal{N}(i)$ representing the neighborhood of i). By substitution in (6), this gives:

$$f_{\text{NW}}(\mathbf{I}_i) = \int Y \frac{\frac{1}{K}\sum_{j\in\mathcal{N}(i)} G_h(\mathbf{I}_i, \mathbf{I}_j) G_g(Y, Y_j)}{\frac{1}{K}\sum_{j\in\mathcal{N}(i)} G_h(\mathbf{I}_i, \mathbf{I}_j)} dY. \tag{7}$$

By assuming G is symmetric about the origin and generalize the expectation such that the observations $\mathcal{I}_k$ are defined in terms of a metric $d_{\mathbb{M}}$ in manifold space d, we obtain:

$$f_{\text{NW}}(\mathbf{I}_i) = \underset{\mathbf{Y}}{\text{argmin}} \frac{\sum_{j\in\mathcal{N}(i)} G(\mathbf{I}_i, \mathbf{I}_j) d_{\mathbb{M}}(Y, Y_j)}{\sum_{j\in\mathcal{N}(i)} G(\mathbf{I}_i, \mathbf{I}_j)} \tag{8}$$

which integrates the distance $d_{\mathbb{M}}(Y, Y_j)$ defined below in Eq.(9) and updates $f_{\text{NW}}(\mathbf{I}_i)$ using the closest neighbors of point $\mathbf{I}_i$ in $\Re^D$. This constrains the regression to be valid for similar data points in its vicinity since locality around $\mathbf{I}_i$ preserves locality in Y.

Using the smooth function $f_{\text{NW}} : \Re^D \rightarrow \mathcal{M}_k$, a new manifold coordinate Y_{new} is created such that $f_{\text{NW}}(\mathbf{I}_{new}) \rightarrow Y_{new}$, based on it's high-dimensional representation. From the intrinsic coordinates of Y_{new}, we approximate it's deviation from the healthy distribution by calculating the geodesic distance based on the shortest path to the tangent plane of the manifold medoid point Y_{med} (center of the manifold). We achieve this by summing a series of pairwise distances between n locally linear patches Ω_i, linking the new data point in $\mathcal{M}_k$ to Y_{med}, with $Y_{new} \in \Omega_0$ and $Y_{med} \in \Omega_{n-1}$. For a 2-chain patch (i and it's neighboring patch j), the point Y_{ij} belonging to the edge shared by Ω_i and Ω_j is determined via convex programming [9]. Generalizing along the entire path, the estimated distance gives:

$$d_{\mathbb{M}}(Y_{new}, Y_{med}) = \sum_{i=0}^{n-1} \min_{Y_{ij}} \|Y_i - Y_{ij}\| + \|Y_j - Y_{ij}\|, \tag{9}$$

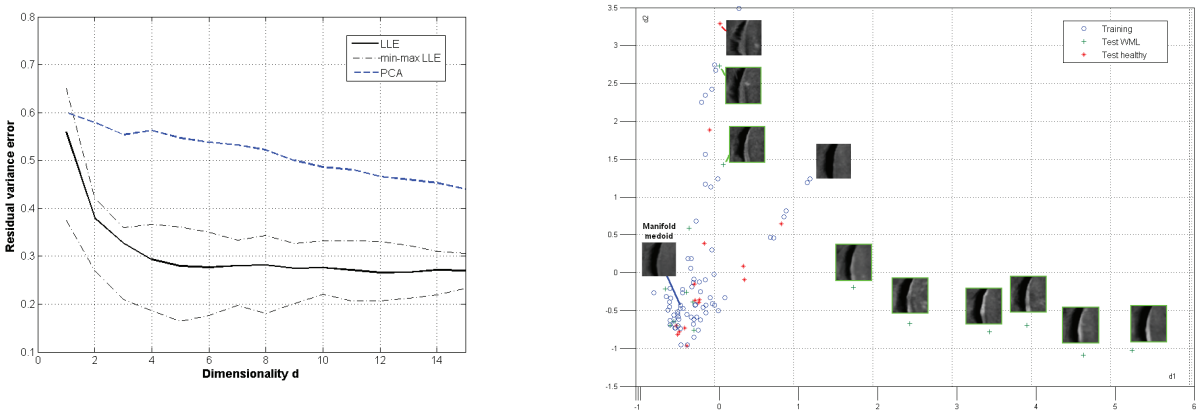
where the estimated geodesic distance $d_{\mathbb{M}}$ is calculated from the low-dimensional mapping and $\{Y_i \in \Omega_i, Y_j \in \Omega_j, Y_{ij} \in (\Omega_i \cap \Omega_j)\}$ are vertices along the manifold path from Y_{new} to Y_{med}.

3. EXPERIMENTS AND RESULTS

Data images used in this study were obtained from a diabetic population of 106 elderly patients. The patient mean age was 60 ± 4.7 (range $54 - 77$, median 59). Forty-nine were female and 57 were male. MRIs were performed during the baseline period on enrollment into the study (TR = 8000 TR, TE = 100). All participant exams consisted of transaxial T1-w, T2-w, PD, and FLAIR scans. All scans except T1-w were performed with a 3-mm slice thickness, no slice gap, a 240×240mm field of view and a 256×256 matrix. From these brain images, 12 periventricular locations were identified and for each of these locations, 73 healthy training and 33 test image patches (39×39) were extracted. Ground-truth lesions were manually segmented by an expert rater on the test images and classified as healthy or with WMLs according to the total lesion load. From the 396 patches, 263 were classified as healthy (no lesion load), and 133 with WMLs. All images were co-registered to a template using HAMMER [10].

The training patches were embedded in the landmark specific-manifold (12 separate manifolds for each of the 12 landmarks) using the approach in 2.1. Then for each test image patch, we would determine the corresponding manifold $\mathcal{M}_k$ based on the weighted similarity measure $d_{\mathbb{A}}(\mathbf{I}_t, \mathbf{I}_{med})$ in 2.1, where $\mathbf{I}_t$ is the test image and $\mathbf{I}_{med}$ is the training image corresponding to manifold medoid Y_{med} in $\mathcal{M}_k$. The test image is then mapped onto the selected manifold to compute the patchwise geodesic distance to manifold center Y_{med}. We tested the manifold selection process used to map each test image patch into the appropriate manifold. Fig. 2 illustrates some results to identify the correct training dataset for a series of test images.

The optimal neighborhood size was found at $K = 10$ for all manifolds in the training set. The optimal manifold dimensionality was estimated by computing the residual reconstruction error between the ambient and intrinsic spaces for each periventricular population, as shown in Fig. 3a. The manifold dimensionality range was $d = [4, 8]$ depending on the selected training set. One can observe the non-linear trend of the curve when comparing the evolution to a linear approximation of the manifold based on PCA. A 2D embedding of the computed manifold (output of LLE) is represented in Fig. 3b, showing the group of training healthy patients in the left portion of the 2D map and the test images which either fall onto the normal tissue manifold or deviate from the manifold in an orthogonal fashion. The bandwidths for the kernels G_h and G_g were set as average distances over the local neighborhoods in ambient and manifold space, respectively. Fig. 4 represents the projected distances between all the images (training and testing) involved in this study and the medoids for each of the 12 manifolds. Images of healthy tissue obtain lower geodesic distances, while most of the images with WMLs demonstrate larger values spanning outside the training cluster. By defining a manifold-specific threshold based on the furthest



(a) (b)

Fig. 3. (a) Evolution of the residual variance error against dimensions d, comparing PCA and LLE. (b) 2D LLE manifold embedding of images from the parietal lobe and corpus callosum areas, according to its two first dimensions (d_1, d_2). Blue points correspond to the training images of healthy tissue, while green crosses and red stars represent testing images for WMLs and healthy tissue respectively.

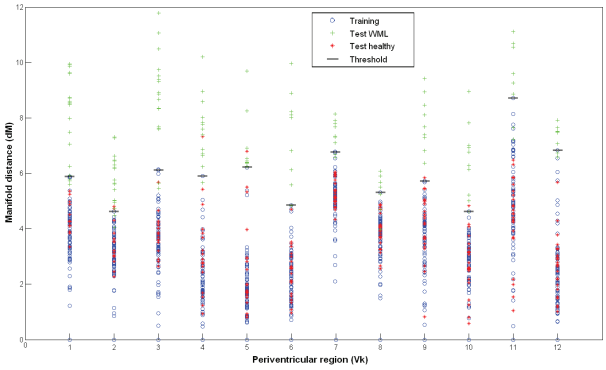


Fig. 4. Distribution of geodesic distances for all images in the study, assessing the presence of WMLs by a deviation to the manifold.

healthy point to identify the presence of WMLs in the image, the total WML detection rate was of 85.7% from all 396 test images, with a false positive rate of 2.2%. The 14.3% of missed cases represents test images with small lesion loads and may vary based on the selected threshold.

4. CONCLUSION

We have proposed a method for detecting the presence of white matter lesions in FLAIR images using a manifold-constrained embedding approach. The set of training manifolds from 12 ventricle locations represents the distribution of healthy images in high dimensional space. A weighted graph is implemented as a non-local similarity measure to select local neighborhoods in the datasets, while a Gaussian regression kernel is used to map unseen image patches onto the training manifold created by LLE. A chain of locally linear patches in manifold space is applied to assess the deviation from normality. Experiments demonstrate the need of nonlinear embedding of the learning data, and the relevance of the proposed method for stratifying different stages of WML progression. In the context of multiple sclerosis, the method can improve for the early detection of the disease with promising classification rates based on ground-truth knowledge. Future work will improve the deviation threshold selection by cross-validation and investigate into fully automated WML segmentation for monitoring and therapy applications.

5. REFERENCES

[1] D. Mortazavi, AZ. Kouzani, and H. Soltanian-Zadeh, "Segmentation of multiple sclerosis lesions in MR images: a review," *Neuroradiology*, 2011.

[2] K.J. Worsley, J.E. Taylor, F. Tomaiuolo, and J. Lerch, "Unified univariate and multivariate random field theory," *NeuroImage*, vol. 23, pp. S189–S195, 2004.

[3] M. Duchateau, M. De Craene, G. Piella, and A.F. Frangi, "Characterizing pathological deviations from normality using constrained manifold-learning," in *MICCAI*, 2011, pp. 256–63.

[4] C. Wachinger, M. Yigitsoy, and N. Navab, "Manifold learning for image-based breathing gating with application to 4D ultrasound," in *MICCAI*, 2010, vol. II, pp. 26–3.

[5] S. Gerber, T. Tasdizen, P. Fletcher, S. Joshi, and R. Whitaker, "Manifold modeling for brain population analysis," *Med. Image Anal.*, vol. 14, pp. 643–653, 2010.

[6] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Sci.*, vol. 290, pp. 2323–6, 2000.

[7] F. Rousseau, P. Habas, and C. Studholme, "Human brain labeling using image similarities," in *CVPR*, 2011, pp. 1081–8.

[8] B. Davis, P.Fletcher, E.Bullitt, and S. Joshi, "Population shape regression from random design data," *Int. J. Comput. Vis.*, vol. 90, pp. 255–266, 2010.

[9] D. Meng, Y. Leung, Z. Xu, T. Fung, and Q. Zhang, "Improving geodesic distance estimation based on locally linear assumption," *Patt. Reco. Letters*, vol. 29, pp. 862–871, 2008.

[10] D. Shen and C. Davatzikos, "HAMMER: hierarchical attribute matching mechanism for elastic registration," *IEEE TMI*, vol. 21, pp. 1421–39, 2002.