Published in final edited form as:

IEEE Trans Med Imaging. 2014 June ; 33(6): 1236–1247. doi:10.1109/TMI.2014.2305751.

Regional Manifold Learning for Disease Classification

Dong Hye Ye[*],
Department of Electrical and Computer Engineering, Purdue University, Lafayette, IN 47907 USA

Benoit Desjardins,
Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104 USA

Jihun Hamm,
Department of Computer Science and Engineering, Ohio State University, Columbus, OH 43210 USA

Harold Litt, and
Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104 USA

Kilian M. Pohl
Center for Health Sciences, SRI International, Menlo Park, CA 94025 USA, and also with the Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA 94304 USA

Abstract

While manifold learning from images itself has become widely used in medical image analysis, the accuracy of existing implementations suffers from viewing each image as a single data point. To address this issue, we parcellate images into regions and then separately learn the manifold for each region. We use the regional manifolds as low-dimensional descriptors of high-dimensional morphological image features, which are then fed into a classifier to identify regions affected by disease. We produce a single ensemble decision for each scan by the weighted combination of these regional classification results. Each weight is determined by the regional accuracy of detecting the disease. When applied to cardiac magnetic resonance imaging of 50 normal controls and 50 patients with reconstructive surgery of Tetralogy of Fallot, our method achieves significantly better classification accuracy than approaches learning a single manifold across the entire image domain.

Index Terms

Abnormality detection; cardiac magnetic resonance imaging (MRI); manifold learning; morphological classification; tetralogy of Fallot (TOF)

I. Introduction

Identifying and classifying the impact of a disease from high-dimensional medical images generally relies on accurately describing the shape of the displayed anatomy. For example, clinicians measure the volume of the ventricles in cardiac magnetic resonance images (MRI)

[*] yed@purdue.edu.

to assess the impact of reconstructive surgery for Tetralogy of Fallot (TOF) [1], [2]. To improve the timing of follow-up surgery (prosthetic pulmonary valve placement), researchers are seeking more refined measurements as, the volumetric scores are insensitive towards the subtle residual impact of the reconstructive surgery [3]–[8]. Many morphometric group studies refine volumetric measurements via deformation-based encodings [9]–[14], which quantify shapes relative to each location in the image domain. The encoding enables these studies to investigate the entire image domain for the impact of disease without prior domain knowledge, i.e., the analysis is performed without experts initially narrowing down the search to specific regions. Using these high-dimensional (on the order of the number of image voxels) encodings, automatically detecting a disease in individual subjects is difficult. It requires complex classifiers, such as [15], [16], to identify the few disease-impacted measurements, e.g., those of a specific anatomical region [17], [18]. To avoid over-fitting, these classifiers require many more samples than provided by typical medical imaging studies. In this paper, we instead reduce the encoding to a set of compact descriptors, which are sensitive towards the disease of interest according to the samples of the imaging study. We do so by introducing the concept of *regional manifold learning*. We then identify the local impact of disease in an individual image by training a simple classifier [19] for each descriptor and classify the image by fusing those results to a single label.

Manifold learning as the means of reducing the complexity of high dimensional descriptors has been applied to a wide variety of medical imaging applications. For example, linear methods, such as principal component analysis (PCA) and independent component analysis (ICA), are used for image segmentation [22], [23], shape analysis [24], and the registration of perfusion images [25]. A limiting factor of the resulting linear manifold is the underlying assumption of the image data being on, or close to, a linear subspace, which is often not the case in the medical domain. Nonlinear manifold learning techniques, such as [20], [26], [27], overcome these limitations by assuming that a set of images lies on a sub-manifold of much lower dimension. They determine the sub-manifold by first arranging the images according to a graph. Each vertex of the graph represents a scan and the weighted edges represent similarity/dissimilarity between the scans. This graph is then embedded into a low-dimensional Euclidean space, preserving the relationship between images as defined by the graph. Thus, if one constructs the graph by measuring shape differences between images, then the coordinates of the resulting low-dimensional embedding can be viewed as a compact descriptor of anatomical variations. Unlike other dimensionality reduction methods, such as [28], the coordinates can also reflect specific anatomical properties and the continuous progression from normal controls to the diseased population.

Examples of nonlinear manifold learning in medical imaging are cardiac segmentation [29], brain registration [30], automatic detection of neurological diseases [31], [32], and the analysis of neonatal brain images [33]. These methods use a single metric to determine the similarity between images. If the metric accounts for anatomical variation across the entire image domain, then disease inflicted changes confined to small regions are most likely obscured by the normal variation in the remainder of the image domain. In addition, the resulting manifold is generally an inaccurate approximation of the true data structure, as the dense-sampling assumption of manifold learning requires a very large number of images, which typical medical studies lack [34]. Most methods, which use manifold learning as a

compact descriptor of high-dimensional encodings, thus rely on an expert to confine the manifold to the image region most likely impacted by disease [35]. To avoid expert intervention, recent publications propose to learn manifolds of multiple image regions and combine their results [36]–[38]. For example, Bhatia *et al.* [39] divide an image into regular patches to learn the manifold of the cardiac and respiratory cycle associated with each patch. These methods require the manifold of a specific patch to be consistent with the manifold trained on the entire image domain. Due to the issues related to learning the manifold of the image domain, these consistency constraints most likely reduce the accuracy of the patch-specific manifolds in capturing the local impact of a disease. In this paper, we avoid this constraint by relying on a bottom-up approach: we separately learn the manifolds for each image region, which we call *regional manifold*, and then perform disease classification by combining their results. Relatively small regions also contain small anatomical variation. Our regional manifolds thus only require a small training set of normal and diseased subjects to accurately capture the shape variations within their corresponding regions.

As outlined in Fig. 1 and Section II, our bottom-up approach automatically identifies cardiac disease, such as TOF, from MRIs by applying a deformation-based descriptor to each region, extracting the disease-relevant information from the descriptor, and finally classifying first each region and then the entire image. Deformation-based descriptors targeted towards describing the anatomy of small image regions, such as RAVENS [14], separately segment the structure in each training case and then nonlinearly register the training case to a template. The resulting deformation encodes the expansion or contraction of the anatomy within this region with respect to the template. In order for the descriptor to be meaningful, the segmentations need to be consistent and accurate across the patient population. While one can parcellate the complex shape of the gray matter in the brain into a set of small regions, the same task is extremely difficult for the ventricles of the heart due to their smooth and simple shapes. Our shape encoding, called inverse RAVENS or *i*RAVENS, circumvents this issue by only parcellating the template into regions of interest (ROI). Similar to atlas-based segmenters [40] but unlike existing deformation-based encodings, *i*RAVENS then nonrigidly registers the template to the subject's scan. It uses the resulting deformation map to describe the ROI-specific anatomy of a subject but unlike existing deformation based encodings for small anatomical regions, such as [14], does not require a detailed segmentation of each scan, is not impacted by the variability introduced by inter-subject segmentation, and allows us to freely choose the definition of the ROIs as they are not confined by visible landmarks in the image. All of these benefits of *i*RAVENS are important aspects in identifying and classifying disease based on regional manifold learning. However, one can easily replace *i*RAVENS with many other encodings of local shape characteristics.

For each ROI, we then derive a disease sensitive descriptor by first mapping the *i*RAVENS maps of the subject to the lower-dimensional embeddings of the regional manifolds, which were created via ISOMAP [20]. Unlike other nonlinear manifold learning techniques, such as [26], [34], ISOMAP preserves the original distance measured across all training image samples when positioning those samples in the lower-dimensional coordinate system. This coordinate system thus protects the anatomical variation captured by the distance score. To transform the compact encoding into a disease sensitive one, we carefully select embedding

coordinates. Most medical image analysis based on manifold learning only keeps the first n embedding coordinates [31], [41], [42] for some fixed n. This strategy ignores the relative importance of individual coordinates for identifying the disease so that some of them may be irrelevant for discriminating TOF from normal scans. We instead only keep the most informative embeddings by identifying them during training phase via minimum redundancy maximum relevance (mRMR) algorithm [21]. The algorithm selects the embedding coordinates which have the highest relevance for distinguishing the two groups while minimizing the dependency between selected embedding coordinates.

The compact and disease-sensitive encoding is passed on to a linear support vector machine (SVM) [19], such as also in [39], to assign each ROI of an image to one of the two populations. We then visualize the outcome across all regions, which we call abnormality maps, by computing the weight of each region according to the outcome of the corresponding classifier and the accuracy of that classifier in identifying the disease on the training data set. We color the region according to the weights and map the corresponding image back to the subject space. The final step of our approach described in Section II combines the weights into a single image label via *late fusion* [43]. Unlike *early fusion*, which equally weighs the importance of the regions in the final labelling, our late fusion approach models the impact of diseases on specific regions by weighing the importance of each image region based on the accuracy of the corresponding classifier in identifying the disease during training [44].

Section III summarizes the results of our experiment, in which we measure the accuracy of different classifiers on distinguishing the MRIs of 50 normal controls and 50 TOF patients via ten-fold cross validation. We specifically chose this data set due to the presence of ground-truth diagnosis, i.e., each patient received corrective surgery during infancy or not, and the fact that the right ventricle is abnormal in all TOF patients. On this data set, we first demonstrate that our regional manifolds are more accurate in describing disease related anatomical changes than ventricular volume, the shape descriptor commonly extracted from cardiac MRIs by clinicians. We then show that our labelling based on the regional manifolds is significantly more accurate than classifiers based on manifold learning over the entire image domain. Furthermore, our method identifies specific regions that may be affected by TOF. These are predominantly in the right ventricle but also include parts of the left ventricle, which echoes recent clinical findings [45].

In summary, the main novelty of this paper is the introduction of regional manifold learning based on the bottom-up principle, i.e. our approach separately learns the manifold for each ROI and then combines the result to a single decision without requiring any prior information about the disease. All important parameters of the approach are automatically determined via well-tested criteria or parameter exploration. While we focus this article on the identification of TOF from cardiac MRI, our regional manifold based classifier is not specific to this disease or organ. Thus, we expect our findings related to regional manifold learning to translate to many other medical imaging applications.

II. Regional Manifold-Based Classification

Following the outline of Fig. 1, we now construct our regional manifold based classifier, which labels images without prior disease information. We first introduce *i*RAVENS, the high dimensional shape encoding extracted from each image region of the input MRI (Section II-A).We then train a manifold learning approach to transform *i*RAVENS into a compact and disease-sensitive descriptor (Section II-B). Based on these descriptors, we train a classifier for each region and define the fusion of the results of the classifiers to a single label (Section II-C). We automatically determine the optimal setting of the components described in Sections II-A, II-B, and II-C via parameter exploration and well-tested criteria on the training data T. This training data consist of N cardiac MRI scans such that $T := \{I_i : i = 1, \ldots, N\}$.Our method assumes that each MRI I_i only shows the left ventricle and right ventricle with the remaining parts being blacked out. The purpose of omitting noncardiac structures, which is similar to skull stripping in brain imaging, is to ensure that our approach is solely influenced by cardiac tissue and not by neighboring anatomies, such as the lung and liver. After completing the training phase of our approach, we automatically classify new MRIs according to Section II-D.

A. iRAVENS: A Deformation-Based Encoding

To define our high-dimensional shape descriptor of the image domain, we automatically select a template image I_T from the training set T, parcellate the image domain of I_T into a grid of disjoint ROIs, and then specify the process of extracting the ROI-specific *i*RAVENS maps.

We carefully choose the template I_T, as a bad selection can negatively influence the accuracy of our classification framework. If I_T is very different from the other scans, its ROIs are difficult to match to the other subjects. This lowers the quality of the *i*RAVENS maps, which rely on accurate nonrigid registration. We reduce this risk by performing unbiased template selection via GRAM [30]. GRAM selects as template, called I_T, the scan that is closest to the geodesic mean [46] of all scans in T. I_T is thus the scan most similar, and likely easiest to register, to all the other scans in the training set.

After selecting the template I_T, we linearly align each image $I_i \in T$ to I_T. Furthermore, we separate the domain Ω of I_T into R disjoint ROIs

$$\Psi := \{\Omega^r \subset \Omega : r = 1, \ldots, R | \Omega^i \cap \Omega^j = \varnothing \text{ for } i \neq j\}.$$

Keeping the regions disjoint, such as in the examples of Fig. 2, allows our framework to independently analyze each region for impact of disease.

Inspired by RAVENS maps [14] and atlas-based segmentation [40], we compute the *i*RAVENS map of the image $I_i \in T$ with respect to Ω^r by first registering I_T to I_i via the diffeomorphic demons algorithm [47]. This results in the deformation map $\phi : x \rightarrow x + \phi_i(x)$ at voxel $x \in \Omega$. Now let $J(\phi_i)$ be the Jacobian determinant of ϕ_i, B^r be the binary map representing the region Ω^r on the image domain Ω, and $B^r(\phi_i)$ the template region warped to

the corresponding subject region. The *i*RAVENS map D_i^r represents the ROI Ω^r with respect to image I_i by the product between $B^r(\phi_i)$, which encodes the boundary of the aligned ROI, and $J(\phi_i)$, which encodes the changes in the shape between the original and aligned ROI with respect to each image location. At voxel $x \in \Omega$, the *i*RAVENS map is thus defined as

$$D_i^r(x) := J(\phi_i(x))\, B^r(\phi_i(x)).$$

Fig. 3(c) is an example of an *i*RAVENS map. The map was created by registering the template of Fig. 3(a), which is a MRI of a normal control, to the TOF subject in Fig. 3(b), whose right ventricle is dilated—a hallmark for the long term impact of TOF. The *i*RAVENS map captures the dilation of the right ventricle via local expansion (red) and by a ROI that has increased in size compared to its origin in the template space. Note that the boundary of the ROI in *i*RAVENS does not perfectly match the right ventricle in Fig. 3(b) as the underlying deformation ϕ_i is too stiff. While a more flexible deformation could perfectly match the boundaries, the corresponding *i*RAVENS map would be more susceptible to noise, which could negatively impact further analysis. Unlike RAVENS, who only encodes local shape information as the ROI of each subject is analyzed in the template space, *i*RAVENS approximates the true boundary of the ROI in the subject space, which provides important information for identifying disease inflicted changes in anatomy as our experiments of Section III indicate.

In summary, we compute the *i*RAVENS map for a specific I_i based on the Jacobian of the deformation from the template to I_i. Repeating this process for each image in T results in the set of *i*RAVENS maps $\{ D_1^r \ldots D_N^r \}$.

B. Creating Compact and Disease-Sensitive Encodings

For each ROI Ω^r, we now derive a compact and disease-sensitive encoding from the *i*RAVENS map. We first learn the regional manifold across the training set T, then use the resulting low-dimensional coordinate system to generate a compact encoding of the *i*RAVENS maps. Finally, we transform this compact representation into a disease-sensitive one by only keeping those coordinates that are most informative with respect to distinguishing the normal controls from the diseased population.

We create the regional manifolds via ISOMAP [20], which represents the manifold as a graph G^r. Each node of G^r denotes the ROI Ω^r of an image, and the edge length between nodes represent the similarity between the corresponding image patches. ISOMAP then maps G^r into a low-dimensional Euclidean space so that the graph structure is locally preserved.

We construct G^r based on the pairwise distance $d^r(I_i, I_j)$ between I_i and I_j with respect to Ω^r. We choose for $d^r(I_i, I_j)$ the L_2 norm (Euclidean distance) of the corresponding *i*RAVENS maps D_i^r and D_j^r

$$d^r(I_i, I_j) := \sum_{x \in \Omega^r} \left(D_i^r(x) - D_j^r(x) \right)^2. \quad (1)$$

The L_2-distance between *i*RAVENS maps Ω^r confined to measures the structural difference between a pair of scans in that region. Two scans with similar shapes in the ROI have similar *i*RAVENS maps and the distance will be small. The opposite is true for the ROIs of two images with different anatomical shapes (see Fig. 4).

After computing d^r for all pairs of images in T, we build the graph G^r by first assigning the edge length between two nodes *a* and *b* as the distance $d^r(a, b)$. We then apply the *k* nearest neighbor (*k*NN) algorithm [48], in which only *k* edges of each node are retained. According to [30], we set the neighborhood size *k* to the smallest values so that G^r is a connected graph. ISOMAP finds the low-dimensional embedding of this graph by first computing the geodesic distance for all pairs of nodes, which is the sum of the length of the edges along the shortest path between the two corresponding images. The results are stored in a matrix with the size corresponding to the number of images. The n^r-dimensional embeddings are the results of applying multi-dimensional scaling (MDS) [49] on the distance matrix. As suggested in [30], n^r is automatically set as the smallest number of embedding coordinates that conserve at least 90% of the original information, which we approximate by the *normalized compactness* [50], i.e., the sum of the eigenvalues associated with the first n^r-dimensional embedding coordinates divided by the sum of all eigenvalues as computed by MDS. Thus, the region Ω_r of training image $I_i \in$ T is now reduced to an n^r-dimensional vector $E_i^r := (e_{i,1}^r, \ldots, e_{i,n^r}^r)$ of the embedding coordinates. Consequently, $\mathrm{E}^r := \{E_1^r, \ldots, E_N^r\}$ are the compact encodings of all training images in T. Fig. 5 shows an example of the 2-D embedding of the right ventricle region. Red and green points represent images of TOF patients and normal controls, respectively. The blue lines are the edges between *k*-nearest neighbors, which connects images with similar right ventricular shapes. Interestingly, the first embedding coordinate captures right ventricular dilation with its tendency increasing for images further to the right. In addition, the graph is nicely separating most of the TOF subjects from the normal controls.

To turn the compact encoding into a disease-sensitive one, we now only keep the most informative embeddings with respect to distinguishing the normal controls from the diseased population by applying minimum redundancy maximum relevance (mRMR) [21] to the coordinates of E^r. Specifically, let $\mathbf{y} := \{y_i \in \{0, 1\} : i = 1, \ldots, N\}$ be the set of labels associated with training data set, i.e., $y_i = 1$ if image I_i is of a TOF patient and $y_i = 0$ if it is normal. Given E^r and **y**, mRMR selects the m^r-dimensional embedding coordinates $\hat{\mathrm{E}}^r \subset \mathrm{E}^r$ that minimize the mutual information between selected embedding coordinates $\hat{\mathrm{E}}^r$ (redundancy) and maximize the average mutual information between $\hat{\mathrm{E}}^r$ and **y** (relevance). In other words, mRMR selects the embedding coordinates with highest relevance to the disease labels **y** while minimizing the dependency between selected embedding coordinates. The corresponding indices of the selected coordinates are denoted as $S^r := \{s_j^r | 1 \leq s_j^r \leq n : j = 1, \ldots, m\}$.

The procedure for computing E^r, $\hat{E}^r$, and S^r is repeated independently for all regions $\Omega^r \in \Psi$, which results in $\{\hat{E}^1, \ldots, \hat{E}^R\}$ and $\{S^1, \ldots, S^R\}$.

C. Training and Fusing Regional Classifiers

We identify the possible impact of a disease by training a classifier for each region and assign a single label to the input MRI by combing the output of the regional classifiers according to their accuracy in identifying the disease.

For each region Ω^r, we train a linear support vector machine (SVM) [19] based on the disease-sensitive, regional embeddings $\hat{E}^r$ and the set of labels $\mathbf{y}$. During training, SVM determines a linear hyperplane that separates the normal and diseased images. We repeat this process for each ROI Ω^r resulting in the set of regional classifiers, which identify the possibly impact of TOF in individual regions.

To derive a single label $y \in \{0, 1\}$ for a (training) image, we combine the results of the region-specific classification by weighing them according to importance of the corresponding region in identifying the disease. We define the importance of each regional classifier based on their confusion matrix during training, i.e., true positive (TP^r), false negative (FN^r), false positive (FP^r), and true negative (TN^r). Specifically, we compute the training accuracy

$$w^r := \frac{TP^r + TN^r}{(TP^r + FP^r + TN^r + FN^r)} \quad (2)$$

which we view as an indicator for the impact of a disease on the region and thus the importance of the region in distinguishing a normal control from a diseased subject.

Fig. 6 visualizes the "w^r" accuracy maps of 100 subjects with respect to different image parcellations. In (B)–(F), regions of high importance (85%, orange or red) are near the septum (the wall separating left ventricle and right ventricle) area which are implicated by reconstructive TOF surgery [51], [52]. They also show regions of importance inside the left ventricle, which echoes recent clinical findings [45]. However, most regions in the left ventricle provide low accuracy, which may explain why volumetric image studies often only identify the impact of TOF in the right ventricle. Furthermore, most regional classifiers on the small grid shown in (G) are of low accuracy (blue) indicating that the size of each region is too small to reliably identify the disease. On the other hand, the regions of the coarse grid in (A) are too large, as none of the corresponding classifier are very accurate (red) in identifying the disease.

Based on the regional weights w^r, we now turn the regional labels $z := (z^1, \ldots, z^R)$of a specific image into a single label y via weighted majority voting, i.e., we compute the normalized sum across the regional weight scores

$$w_v(z) := \frac{1}{\sum_{r=1}^{R} w^r} \sum_{r=1}^{R} w^r \cdot z^r$$

then label the image as diseased $w_v(z) > 0.5$ if and normal otherwise.

The product rule [53] is an alternative to fusing regional classifiers based on the confusion matrix. Unlike weighted majority voting, the product rule weights the importance of a regional classifier by separately computing a weight for each output of the classifier, i.e., {0,1}. In case the classifier of region Ωr was positive, $z^r = 1$, then we weight region Ω^r by the true- and false-positive values

$$w_r(1) := \frac{\mathrm{TP}^r}{\mathrm{TP}^r + \mathrm{FP}^r}.$$

In case it was not positive, $z^r = 0$, then we weight the corresponding region according to the true and false negative values

$$w_r(0) := \frac{\mathrm{FN}^r}{\mathrm{FN}^r + \mathrm{TN}^r}.$$

Next, the product rule computes a weight for each possible image label y by the product across all the regional weights, i.e.,

$$w_p(y, z) := \prod_{r=1}^{R} w_r(z^r)^y \left(1 - w_r(z^r)\right)^{1-y}. \quad (3)$$

Note that if the regional label z^r agrees with the global label y, then the component is $w_r(z^r)^y (1 - w_r(z^r)^y)$ the normalized "true" value, i.e., $w_r(0)^0 \cdot (1 - w_r(0))^1 = 1 - w_r(0) = \mathrm{TN}^r/(\mathrm{FN}^r + \mathrm{TN}^r)$. Finally, the product rule assigns the image to the label with the maximum weight, i.e., the image is labelled as diseased if $w_p(1, z) > w_p(0, z)$ and normal otherwise. We expect the product rule to weight individual regions similar to the weighted majority Vote as both methods compute the weights based on the confusion matrix. However, the product rule infers the image label y by taking into account the sensitivity and specificity of each regional classifier. This is not the case for the weighted majority voting, which simply ignores regional classifiers that are not accurate in identifying the disease.

Having defined the mechanism for assigning a single label to an image completes the training of our approach, which was done without any prior information about the disease.

D. Detecting Disease in New Scans

Our approach now classifies a new (test) image I_t, that is not part of the training set T, by following the workflow of Fig. 1, whose steps we now review for test cases.

After computing the *r*RAVENS maps for each region, our method determines the regional embedding E_t^r for I_t by first computing the distances $d^r(I_t, I_i)$ of I_t to the training images $I_i \in$ T [see (1)]. Based on the distances $d^r(I_t, \cdot)$ I_t, is projected to the corresponding regional manifold via Landmark-ISOMAP [54]. This method first adds I_t to the existing *k*NN graph of the regional manifold by adding edges between I_t and the k most similar images in T according to $d^r(I_t, \cdot)$. Then, it computes the geodesic distances for I_t to all the training images and derives the regional embedding coordinates $E_t^r := (e_{t,1}^r, \ldots, e_{t,n}^r)$ for I_t. This approach thus preserves the regional embeddings E^r for the images of T. Next, only those coordinates of $\hat{E}_t^r$ are kept that are listed in S^r, which results in the compact and disease-sensitive embedding $\hat{E}_t^r := (e_{t,s_1^r}^r, \ldots, e_{t,s_m^r}^r)$. Finally, $\hat{E}_t^r$ is fed into the corresponding regional classifiers. The process is repeated for all R ROIs, resulting in the set of regional labels $\mathbf{z_t} := \{z_t^r \in \{0,1\} : r = 1, \ldots, R\}$.

To visualize the possible impact of disease on the individual subject, we multiply these labels $\mathbf{z_t}$ (which we now assume to be {−1, +1}instead of {0,1}) with the training accuracy scores of the regional classifiers of (2), color the regions according to these values, and map the corresponding chart from the template to the subject space. Fig. 7 shows the resulting abnormality maps with grid size 15 × 15 × 8 mm [as in Fig. 6(d)] of three normal subjects and three TOF patients. The maps of the normal subjects misclassify a few regions as impacted by disease (red). However, they do so with relatively low confidence as the accuracy of the corresponding classifiers was low during training. In addition, most of the other regions clearly indicate that the subject is normal. The maps of the TOF patients are strikingly different: most of the regions indicate the presence of disease with high weights.

The final step of our approach reduces the abnormality map of image I_t into a single label according to the fusion approaches of Section II-C.

III. Classification Accuracy

We measure the accuracy of our regional manifold-based classifier on a data set consisting of cardiac MRIs of 50 healthy subjects and 50 TOF patients. We use ten-fold cross validation, i.e., we repeatedly train the classifiers on 90 subjects and measure on the remaining 10 subjects the accuracy of the classifier in distinguishing normal controls versus TOF patients. Our results on this data will show that regional manifolds are substantially more accurate in describing disease-related anatomical changes than conventional clinical measures. In addition, our labelling based on regional manifolds outperforms classification based on global manifold learning and detects abnormal regions impacted by TOF. The remainder of this section describes the experimental set up, reports the accuracy scores of our regional manifold-based classifier compared with current methods, and ends with visual inspection of our classification results.

A. Experimental Setup

1) Data Set—Our data set consists of short-axis cine MR images at end-diastole of 50 healthy adults and 50 TOF patients. All cases were selected based on their medical history. Healthy subjects were scanned for diagnostic purpose and confirmed as normal. In addition, their medical records did not list any other cardiac diseases. All TOF patients were 18 years or older, received reconstructive surgery during infancy, and were scanned as part of routine follow-up. In addition, we insured that the two groups were age-matched (normal group: 35.24 ± 13.02 years; TOF group: 32.64 ± 8.91 years, two-sided unpaired t-test: $p > 0.1$).

Each scan was acquired as part of standard clinical care at the Hospital of the University of Pennsylvania. The acquisition consisted of a sequence of balanced steady state free precession (SSFP) short-axis images of the heart using breath holds in a 1.5T Siemens Avanto scanner with 1.25 mm in-plane resolution and 8 mm slice thickness. During preprocessing, each scan was anonymized, and we confined the spatial positions of these short-axis scans to the mid-portion of the ventricles (12 slices), where the pathologies are usually prominent.

An expert segmented the left ventricle and right ventricle using the software “SEGMENT” [55] and manually edited the results for quality control. We then used the segmentation to omit noncardiac structures from the scans. Furthermore, we corrected for slice misalignment due to breathing motion by detecting the center of LV via Hough transform [56] and then stacking the slices so that the center of the left ventricle aligns across the slices. Finally, each scan was bias field corrected via N4ITK [57] and linearly aligned to the Auckland Cardiac Atlas [58] using FSL flirt [59].

2) Parameter Exploration—There are four important parameters in our regional manifold based-classifier, which were not automatically determined in Section II. To begin with, Section II-A did not specify a criterion for the optimal ROI setting when partitioning the image domain. For each region Ω^r, we also have to determine the smoothness parameter for Demons (the registration approach used by the deformation-based encoding *i*RAVENS in Section II-A), the optimal number of embedding coordinates generated by mRMR (Section II-B), and the penalty parameter of the linear SVM classifier (Section II-C). We determine the optimal values for these four parameters by simultaneously exploring their search spaces based on five-fold cross validation on the training data set. We now define the parameters and their search spaces in further detail.

The partitioning scheme of Fig. 2(b) parcellates the image domain into a regular grid. The definition of the grid depends on the grid size “g,” whose search space 75 × 75 × 32 mm (A), 37.5 × 37.5 × 16 mm (B), 18.75 × 18.75 × 8 mm (C), 15 × 15 × 8 mm (D), 12.5 × 12.5 × 8 mm (E), 8.75 × 8.75 × 8 mm (F), 6.25 × 6.25 × 8 mm (G) covers a wide range of sizes as visible in Fig. 6. The optimal grid setting chosen by the parameter exploration switched between (D) and (E) depending on the experiment. The regions are confined to a regular grid to keep the cost low for determining the optimal parcellation, which, for example, is not the case for the irregular mesh proposed in [60].We do, however, also test the partitioning scheme of Fig. 2(a), for which an expert manually segmented the inter-ventricular septum (separating the right from the left ventricle) in the template. Unlike this expert-defined

partitioning scheme, the generic grid-partition does not require any prior knowledge and allows us to display the possible impact of TOF inside the ventricle, such as in Fig. 7.

In order to define the search space of s, the smoothness parameter of the deformation maps of region Ω^r, we first generated the deformation maps for the entire image domain by running Demons with different global smoothing setting s. With respect to s, we then measured the accuracy of the regional manifold-based classifier, which monotonously dropped with increasing s from 92% for $s = 0$ down to 89% for $s = 2.5$. We note that s cannot be negative, and the smaller s the less Demons constrains the flexibility within the deformation map so that the resulting *i*RAVENS maps are better in capturing the subtle anatomical changes caused by TOF. Based on these results, we set the search space for to {0,0.5,1.5} and the optimal smoothness setting chosen by the parameter exploration varied between 0 and 0.5 across all experiments.

For each region Ω^r, we also define the search space for the number of embeddings selected by mRMR, i.e., m^r, with respect to the embedding dimension n^r of ISOMAP. Specifically, we set the search space to $\{0.2n^r, 0.4n^r, 0.6n^r, 0.8n^r, n^r\}$. According to [61], we define the search space of the fourth parameter, the penalty parameter C^r of the linear SVM classifier, as $\{2^{-3}, 2^{-1}, \ldots, 2^7\}$. Across all regions and experiments, the histograms of the optimal setting of those two parameters mirrored hyperbolic functions, whose maxima were at the smallest values and then quickly tapered off.

B. Classification Results

The overall goal of this experiment is to report the accuracy of our classifier in correctly identifying scans as TOF or normal control, investigating the importance of the individual components of our approach, and compare the results with those of current methods. Using the grid partition and combining the regional classifier results via the product rule of Section II-C, the accuracy of our regional manifold-based classification is 92% according to the ten-fold cross validation on the 100 data sets. We note that all important parameters were automatically determined in this experiment. However, the accuracy did not change when fixing the smoothing parameter s^r to zero for all regions (see Section III-A2), so that we did not vary s^r in the remaining experiments.

To further justify the approach of Section II, we compare its accuracy scores to simpler implementations, i.e., our approach by replacing *i*RAVENS with the intensity values of the MRIs (Accuracy: 85%), and by directly applying *i*RAVENS to the linear SVM (Accuracy: 71%) and a nonlinear SVM [62] (Accuracy: 78%). In all three instances, the accuracy scores were well below the 92% achieved by our method. The accuracy of our method also dropped when we omitted mRMR (Accuracy: 89%) or replaced the *late fusion* strategies of Section II-C with the *early fusion* one (Accuracy: 79%), which first concatenates the disease-sensitive encoding and then applies a linear SVM classifier to resulting vector. To further motivate the individual components of our approach, the remainder of this section compares our method to alternative implementations. We first show that the shape changes caused by TOF are better captured by our disease-sensitive encoding based on the regional manifolds of Section II-B than by the standard clinical shape representation (volume of the left and right ventricle). We then discuss the sensitivity of our approach to the grid size and compare

our findings with those achieved by alternative fusion strategies. Finally, we demonstrate the superiority of our regional manifolds compared to the manifold for the entire image domain.

1) Comparison to Clinical Score—The first experiment compares the accuracy of the disease-sensitive encoding based on the expert-defined partitioning of the ventricles [see also Fig. 2(a)] to the corresponding volumetric measurements. Specifically, we measure the accuracy of a linear SVM classifier applied to each encoding just based on the left ventricle (LV) alone, the right ventricle (RV) alone, as well as both together (LV & RV). Classifications based on both structures first concatenate the features extracted from each region into one vector and then label the scan based on that vector (*early fusion*). This deviates from the *late fusion* approach proposed in Section II-C, which first feeds the features into regional classifiers and then combines the results. We did so in order to directly link the accuracy of the classifier with the quality of the features in encoding anatomical changes caused by the disease. In addition, both encodings performed slightly worse when using late fusion as they ignored the information in the left ventricle. Finally, we note that the volumetric measurements for each ventricle require the segmentation of the inter ventricular septum in each individual case. The disease-sensitive encoding, on the other side, requires this segmentation only to be done on the template [see Fig. 2(a)]. Thus, the classification task based on the regional embeddings is slightly more difficult as it relies on less prior information than the volume scores.

Table I summarizes the accuracy scores. For each of the three experiments (classification based on LV, RV, or LV & RV), our disease-sensitive encoding produced higher classification accuracy than the ventricular volumes, i.e., our regional embeddings are more accurate in capturing the shape change induced by TOF. According to the McNemar's test [63], our encoding was significantly better ($\chi^2 > 3.84$)than the volume measurements in case of RV and LV & RV, but not for LV ($\chi^2 < 3.84$). The McNemar's test is a variant of χ^2 test to analyze matched pairs of data. The test counts the number of cases where one of the classifiers is failed but the other succeeded. It then quantifies the statistical difference between the two classifiers by computing the z-score of these counts.

Looking at each ROI, we observe that both methods perform badly (around 50%) when the classification is confined to the LV. This is not the case for the RV, where both types achieve a classification accuracy above 80%. These results agree with the clinical finding that TOF mostly impacts the RV. Interestingly, both feature types achieve the highest accuracy when concatenating the features of LV and RV (volume: 83%; regional embedding: 89%). This implies that LV contains information useful for disease detection but only in combination with RV. Finally, we note that the accuracy of both encodings was lower than our regional manifold-based classifier, whose grid partitioning can detect the impact of disease in regions within each ventricle.

2) Accuracy of Combining Regional Classifiers—As a follow-up to the previous experiment, we now investigate the accuracy of regional manifold-based classifiers with respect to the grid size of the partitioning, i.e., before running our approach we fix the grid size to a value of the search space specified in Section III-A2 and visualized by Fig. 6. We

then compare those accuracy to the ones achieved by our method, replacing the product rule of Section II-C with weighted majority voting and unweighted majority voting.

Fig. 8 summarizes the outcome of the experiment. The x-axis represents the grid size and the y-axis corresponds to the classification accuracy, with the results of the product rule shown in blue, weighted majority voting in green, and majority voting in red. In general, all combination strategies perform poorly with the largest ROI (A). There are two possible explanations. First, the largest regions have more shape variability and require more training samples to accurately learn the manifold. Given the fixed training size, the accuracy of the embeddings thus generally drops with the increase in size of the corresponding ROI. Second, reconstructive surgery generally impacts small regions of the heart, and therefore a large ROI can dilute information in the small region impacted by reconstructive surgery. In addition, we note that the classification accuracy is low for the smallest ROI (G). Its size is too small for the regional manifolds to capture meaningful anatomical information and thus mostly captures noise in the measurements. The remaining ROI settings (B–F) produce good (≥80%) to very high (≥90%) accuracy scores, which is consistent with our findings with respect to the training accuracy of each regional classifier shown in Fig. 6.

Out of the three ensemble strategies, unweighted majority voting generally performed worse. This finding supports our claim that classifiers should account for the fact that TOF impacts the anatomy of the heart differently in specific regions. Of the two weighted combination strategies, the Product Rule generally was better than weighted majority voting, with the peak accuracy of 92% for grid size (E), which is equivalent to the accuracy of the regional manifold-based classifier parametrized automatically. Weighted majority voting simply ignores regional classifiers that are not accurate in identifying the disease. This is not the case for the product rule, which determines the label of a scan by taking into account the sensitivity and specificity of each regional classifier. The difference in performance between the two strategies indicates that even regions with low accuracy are important for identifying TOF, which echoes our findings of Section III-B1.

3) Global Versus Regional Manifold Learning—In the final experiment, we compared our regional manifold learning approach to alternative manifold learning techniques by replacing ISOMAP with the linear approach called principle component analysis (PCA) [64] and learning a single manifold over the entire image domain, which we refer to as global manifold learning. For all experiments, we fix the grid size to (E) in addition to the regional smoothing parameter s^r as it does no impact the accuracy of our approach according to the exploration in Section III-B2.

For both learning methods, the regional approach (ISOMAP: 92%, PCA: 90%) was significantly more accurate than the global one (ISOMAP: 76%, PCA: 74%) according to McNemar's test ($\chi^2 > 3.84$). These results confirm our earlier observation that encodings of larger image regions tend to be insensitive towards the local changes caused by TOF. In addition, ISOMAP produced slightly higher classification accuracy than PCA for both global and regional manifold learning. We therefore conclude that the nonlinear model of ISOMAP is a better fit than the linear model of PCA for capturing the complex shape variations of the hearts in our data set.

C. Visual Inspection

We complement the quantitative comparison above with a qualitative analysis of the results of our regional manifold-based classifier. Fig. 9 shows a sample slice of each misclassified case. We notice that the misclassified TOF cases only show minor residual effects from the reconstructive TOF surgery and thus look similar to the misclassified normal controls. The reconstructive TOF surgery involves expanding the narrowed right-ventricular outflow tract with a pericardial patch, repairing the septal defect, and dilating the narrowed pulmonic valve. This last part of the procedure invariably leaves the patient with regurgitation of the pulmonic valve. Over time, the increased volume load on the right ventricle leads to characteristic shape changes and dilation, which are shown in the misclassified cases as well as in Fig. 10. In Fig. 10, we sorted all cases according to their weight $w_p(1,z)$ defined by the product rule of (3), which is the weight of the image being labelled TOF, and then selected six representative samples from being most likely normal (left) to most likely TOF (right). The figure illustrates the progression of the shape of the right ventricle from being round in normal controls to the different degrees of dilation and irregularity in TOF patients. The results indicate that the overall weight computed by our regional manifold-based classifier reflects the shape abnormalities in the heart.

In summary, these experiments have not only validated our choices for the regional manifold based classifier, but also showed that this approach is more accurate in encoding the local anatomical impact of reconstructive TOF surgery than volumetric measurements and global manifold learning. Without depending on prior knowledge of the disease, the approach not only automatically classifies individual scans but also provides weight scores that can guide visual comparison of TOF cases. The weight scores might aid clinicians in improving the timing for follow-up surgery in individual TOF subjects, which is one of our long-term goals.

IV. Conclusion

Manifold learning is frequently used in medical image analysis. However, the accuracy of existing implementations suffers from viewing each image as a single data point. In this paper, we proposed an alternative to learning the manifold of the entire image domain by separately learning the manifold for each ROI. We then used the regional manifolds to capture changes in anatomy due to disease specific to each region. We trained a classifier for each disease-sensitive encoding allowing us to identify regions possibly affected by the disease even without any prior knowledge about the disease. We produced a single decision for each scan by combining the regional classification based on the accuracy of that region in identifying the disease. We demonstrated the superiority of the method by distinguishing TOF subjects from normal controls. Based on ten-fold cross validation, our method achieved 92% accuracy, which is significantly better than the classification based on ventricular volume measurements or global manifold learning.

Though we focused herein on identifying scans of TOF patients, our classifier based on regional manifolds is not specific to this disease or the heart. It learns region-specific patterns characterizing the disease from features extracted for all voxels of the MRI. Our approach should thus be applicable to other data sets, such as brain MRIs of normal controls

and Alzheimer patients, where large anatomical shape variations are observed within the healthy population and the disease-inflicted anatomical changes are confined to specific, small image regions [65], [66]. Our regional classifier should also be applicable to diseases regionally impacting other cardiac properties, such as the movement of the heart in case of septal flash [35]. Assuming the feature is a good fit for the disease, we expect the classifier to be more accurate when learning the manifold of image regions instead of the entire image domain. In conclusion, classification based on regional manifolds should be beneficial to many other medical image applications.

Acknowledgments

This work was supported in part by the National Institutes of Health (NIH)under Grant K05 AA017168, in part by the National Center for Research Resources under Grant UL1RR024134, and in part by the Institute for Translational Medicine and Therapeutics' (ITMAT) Transdisciplinary Program in Translational Medicine and Therapeutics.

References

1. Morcos P, Vick I, Wesley G, Sahn D, Jerosch-Herold M, Shurman A, Sheehan F. Correlation of right ventricular ejection fraction and tricuspid annular plane systolic excursion in Tetralogy of Fallot by magnetic resonance imaging. Int J Cardiovasc Imag. 2009; 25(3):263–270.
2. Solarz D, Witt S, Glascock B, Jones F, Khoury P, Kimball T. Right ventricular strain rate and strain analysis in patients with repaired Tetralogy of Fallot: Possible interventricular septal compensation. J Am Soc Echocardiogr. 2004; 17(4):338–344. [PubMed: 15044867]
3. Villafañe J, Feinstein JA, Jenkins KJ, Vincent RN, Walsh EP, Dubin AM, Geva T, Towbin JA, Cohen MS, Fraser C, Dearani J, Rosenthal D, Kaufman B, Graham TP. Hot topics in Tetralogy of Fallot. J Am Coll Cardiol. 2013; 23:2155–2166.
4. Wald RM, Haber I, Wald R, Valente AM, Powell AJ, Geva T. Effects of regional dysfunction and late gadolinium enhancement on global right ventricular function and exercise capacity in patients with repaired Tetralogy of Fallot. Circulation. 2009; 119(10):1370–1377. [PubMed: 19255342]
5. Zhang H, Wahle A, Johnson RK, Scholz TD, Sonka M. 4-D cardiac MR image analysis: Left and right ventricular morphology and function. IEEE Trans Med Imag. Feb; 2010 29(2):350–364.
6. Bernardis, E., Konukoglu, E., Ou, Y., Metaxas, DN., Desjardins, B., Pohl, KM. Medical Image Computing and Computer Assisted Intervention (*MICCAI*). Vol. 7511. New York: Springer; Temporal shape analysis via the spectral signature; p. 49-56.Lecture Notes in Computer
7. Mansi, T., Durrleman, S., Bernhardt, B., Sermesant, M., Delingette, H., Voigt, I., Lurz, P., Taylor, A., Blanc, J., Boudjemline, Y., Pennec, X., Ayache, N. Medical Image Computing and Computer Assisted Intervention (MICCAI). Vol. 5761. New York: Springer; 2009. A statistical model of right ventricle in Tetralogy of Fallot for prediction of remodelling and therapy planning; p. 215-221.Lecture Notes in Computer Science
8. Bove T, Vandekerckhove K, Devos D, Panzer J, De Groote K, De Wilde H, De Wolf D, De Backer J, Demulier L, François K. Functional analysis of the anatomical right ventricular components: Should assessment of right ventricular function after repair of Tetralogy of Fallot be refined? Eur J Cardio-Thoracic Surg. 2013:1–7.
9. Ashburner J, Friston KJ. Voxel-based morphometry—The methods. NeuroImage. 2000; 11:805–821. [PubMed: 10860804]
10. Thompson P, Giedd J, Woods R, MacDonald D, Evans A, Toga A. Growth patterns in the developing human brain detected using continuum-mechanical tensor mapping. Nature. 2000; 404:190–193. [PubMed: 10724172]
11. Gaser C, Volz H, Kiebel S, Riehemann S, Sauer H. Detecting structural changes in whole brain based on nonlinear deformations: Application to schizophrenia research. NeuroImage. 1999; 10:107–113. [PubMed: 10417245]

12. Thirion J, Calmon G. Deformation analysis to detect and quantify active lesions in 3D medical image sequences. IEEE Trans Med Imag. May; 1999 18(5):429–441.
13. Leow A, Klunder A, JC, Toga A, Dale A, Bernstein P, Britson M, Gunter J, Ward C, Whitwell J, Borowski B, Fleisher A, Fox N, Harvey D, Kornak J, Schufff N, SC, AG, WM, Thompson P. Longitudinal stability ofMRI for mapping brain change using tensor-based morphometry. NeuroImage. 2006; 31:627–640. [PubMed: 16480900]
14. Shen D, Davatzikos C. Very high-resolution morphometry using mass-preserving deformations and HAMMER elastic registration. NeuroImage. 2003; 18:28–41. [PubMed: 12507441]
15. Li Y, Gong S, Liddell H. Recognising trajectories of facial identities using kernel discriminant analysis. Image Vis Comput. 2003; 21:1077–1086.
16. Vishwanathan SVN, Sun Z, Theera-Ampornpunt N, Varma M. Multiple kernel learning and the SMO algorithm. Adv Neural Inf Process Syst. 2010:3311–3325.
17. Qian Z, Liu Q, Metaxas DN, Axel L. Identifying regional cardiac abnormalities from myocardial strains using nontracking-based strain estimation and spatio-temporal tensor analysis. IEEE Trans Med Imag. Dec; 2011 30(12):2017–2029.
18. Punithakumar K, Ayed IB, Islam A, Goela A, Ross IG, Chong J, Li S. Regional heart motion abnormality detection: An information theoretic approach. Med Image Anal. 2013; 17:311–324. [PubMed: 23375719]
19. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995; 20:273–297.
20. Tenenbaum JB, de Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. Science. 2000; 290:2319–2323. [PubMed: 11125149]
21. Peng H, Long F, Ding C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell. Aug; 2005 27(8):1226–1238. [PubMed: 16119262]
22. Mitchell SC, Bosch JG, Lelieveldt BPF, van der Geest RJ, Reiber JHC, Sonka M. 3-D active appearance models: Segmentation of cardiac MR and ultrasound images. IEEE Trans Med Imag. Sep; 2002 21(9):1167–1178.
23. Sun, W., Çetin, M., Chan, RC., Reddy, VY., Holmvang, G., Chandar, V., Willsky, AS. Information Processing in Medical Imaging (IPMI). Vol. 3565. New York; Springer: 2005. Segmenting and tracking the left ventricle by learning the dynamics in cardiac images," in; p. 553-565.Lecture Notes in Computer Science
24. Pohl K, Fisher J, Bouix S, Shenton M, McCarley R, Grimson W, Kikinis R, Wells W. Using the logarithm of odds to define a vector space on probabilistic atlases. Med Image Anal. 2007; 11(6): 465–477. [PubMed: 17698403]
25. Milles J, van der Geest RJ, Herold MJ, Reiber JHC, Lelieveldt BPF. Fully automated motion correction in first-pass myocardial perfusion MR image sequences. IEEE Trans Med Imag. Nov; 2008 27(11):1611–1621.
26. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. Science. 2000; 290:2323–2326. [PubMed: 11125150]
27. Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computat. 2003; 15:1373–1396.
28. Qu Y, Adam B, Thornquist M, Potter J, Thompson M, Yasui Y, Davis J, Schellhammer P, Cazares L, Clements M, Wright G, Feng Z. Data reduction using a discrete wavelet transform in discriminant analysis of very high dimensionality data. Biometrics. 2003; 59:143–151. [PubMed: 12762451]
29. Zhang QL, Souvenir R, Pless R. On manifold structure of cardiac MRI data: Application to segmentation," in. Comput Vis Pattern Recognit. 2006:1092–1098.
30. Hamm J, Ye DH, Verma R, Davatzikos C. GRAM: A framework for geodesic registration on anatomical manifolds. Med Image Anal. 2010; 14:633–642. [PubMed: 20580597]
31. Gerber S, Tasdizen T, Fletcher PT, Joshi SC, Whitaker RT. Manifold modeling for brain population analysis. Med Image Anal. 2010; 14:643–653. [PubMed: 20579930]
32. Wolz R, Aljabar P, Hajnal JV, Lötjönen J, Rueckert D. Nonlinear dimensionality reduction combining MR imaging with non-imaging information. Med Image Anal. 2012; 16:819–830. [PubMed: 22244037]

33. Aljabar P, Wolz R, Srinivasan L, Counsell SJ, Rutherford MA, Edwards AD, Hajnal JV, Rueckert D. A combined manifold learning analysis of shape and appearance to characterize neonatal brain development. IEEE Trans Med Imag. Dec; 2011 30(12):2072–2086.
34. Coifman R, Lafon S. Diffusion maps. Appl Computat Harmonic Anal. 2006; 21:5–30.
35. Duchateau N, Craene MD, Piella G, Frangi AF. Constrained manifold learning for the characterization of pathological deviations from normality. Med Image Anal. 2012; 16:1532–1549. [PubMed: 22906821]
36. Cardoso, MJ., Wolz, R., Modat, M., Fox, NC., Rueckert, D., Ourselin, S. Medical Image Computing and Computer Assisted Intervention (MICCAI). Vol. 7511. New York: Springer; 2012. Geodesic information flows; p. 262-270.Lecture Notes in Computer Science
37. Ye, DH., Hamm, J., Kwon, D., Davatzikos, C., Pohl, KM. Medical Image Computing and Computer Assisted Intervention (*MICCAI).* Vol. 7512. New York: Springer; 2012. Regional manifold learning for deformable registration of brain MR Images; p. 131-138.Lecture Notes in Computer Science
38. Ye, D., Hamm, J., Desjardins, B., Pohl, K. Medical Image Computing and Computer Assisted Intervention (MICCAI). Vol. 8151. New York: Springer; FLOOR: Fusing locally optimal registrations; p. 195-202.Lecture Notes in Computer Science
39. Bhatia KK, Rao A, Price A, Wolz R, Hajnal J, Rueckert D. Hierarchical manifold learning for regional image analysis. IEEE Trans Med Imag. Feb; 2014 33(2):444–461.
40. Cabezas M, Oliver A, Llado X, Freixenet J, Cuadra MB. A review of atlas-based segmentation for magnetic resonance brain images. Comput Methods Programs Biomed. 2011; 104:158–177.
41. Wolz R, Aljabar P, Hajnal JV, Hammers A, Rueckert D. Alzheimer's Disease Neuroimaging Initiative. LEAP: Learning embeddings for atlas propagation. NeuroImage. 2010; 49:1316–1325. [PubMed: 19815080]
42. Jia H, Wu G, Wang Q, Shen D. ABSORB: Atlas building by self-organized registration and bundling. NeuroImage. 2010; 51:1057–1070. [PubMed: 20226255]
43. Atrey PK, Hossain MA, Saddik AE, Kankanhalli MS. Multimodal fusion for multimedia analysis: A survey. Multimedia Syst. 2010; 16:345–379.
44. Polikar R. Ensemble based systems in decision making. IEEE Circuits Syst Mag. 2006; 6(3):21–45.
45. Spiewak M, Malek L, Petryka J, Mazurkiewicz L, Werys K, Biernacka E, Kowalski M, Hoffman P, Demkow M, Misko J, Ruzyllo W. Repaired Tetralogy of Fallot: Ratio of right ventricular volume to left ventricular volume as a marker of right ventricular dilatation. Radiology. 2012; 265:78–86. [PubMed: 22771877]
46. Joshi S, Davis B, Jomier M, Gerig G. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage. 2004; 23:S151–S160. [PubMed: 15501084]
47. Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage. 2009; 45:S61–S72. [PubMed: 19041946]
48. Miller GL, Teng SH, Thurston W, Vavasis SA. Separators for sphere-packing and nearest neighbor graphs. J ACM. 1997; 44:1–29.
49. Cox, TF., Cox, MA. Multidimensional Scaling. London, U.K: Chapman Hall; 2001.
50. Davies RH, Twining CJ, Cootes TF, Taylor CJ. Building 3-D statistical shape models by direct optimzation. IEEE Trans Med Imag. Apr; 2010 29(4):961–981.
51. Helbing WA, de Roos A. Clinical applications of cardiac magnetic resonance imaging after repair of Tetralogy of Fallot. Pediatric Cardiol. 2000; 21:70–79.
52. Lai WW, Gauvreau K, Rivera ES, Saleeb S, Powell AJ, Geva T. Accuracy of guideline recommendations for two-dimensional quantification of the right ventricle by echocardiography. Int J Cardiovasc Imag. 2008; 24:691–698.
53. Kim H, Ghahramani Z. Bayesian classifier combination. J Mach Learn Res. 2012; 22:619–627.
54. de Silva V, Tenenbaum JB. Global versus local methods in nonlinear dimensionality reduction. Adv Neural Inf Process Syst. 2002:705–712.
55. Segment. [Online]. Available: http://medviso.com/products/segment/

56. Duda R, Hart PE. Use of the Hough transformation to detect lines and curves in pictures. Commun ACM. 1972; 15:11–15.
57. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC. N4ITK: Improved N3 bias correction. IEEE Trans Med Imag. Jun; 2010 29(6):1310–1320.
58. Cardiac atlas. [Online]. Available: http://atlas.scmr.org/
59. Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M, Smith SM. Bayesian analysis of neuroimaging data in FSL. NeuroImage. 2009; 45:S173–186. [PubMed: 19059349]
60. Sundar, H., Shen, D., Biros, G., Xu, C., Davatzikos, C. Medical Image Computing and Computer Assisted Intervention (MICCAI). Vol. 4791. New York: Springer; 2007. Robust computation of mutual information using spatially adaptive meshes; p. 950-958.Lecture Notes in Computer Science
61. A Practical Guide to Support Vector Classification. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf
62. Keerthi S, Lin C. Asymptotic behaviors of support vector machines with Gaussian kernel. Neural Computat. 2003; 15:1667–1689.
63. Dietterich TG. Approximate statistical tests for comparing supervised classification learning algorithms. Neural Computat. 1998; 10:1895–1923.
64. Jolliffe, IT. Principal Component Analysis. NewYork: Springer; 1986.
65. Convit A, de Asis J, de Leon M, Tarshish C, de Santi S, Rusinek H. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimers Disease. Neurobiol Aging. 2000; 21:19–26. [PubMed: 10794844]
66. Fan Y, Shen D, Gur RC, Gur RE, Davatzikos C. COMPARE: Classification of morphological patterns using adaptive regional elements. IEEE Trans Med Imag. Jan; 2007 26(1):93–105.

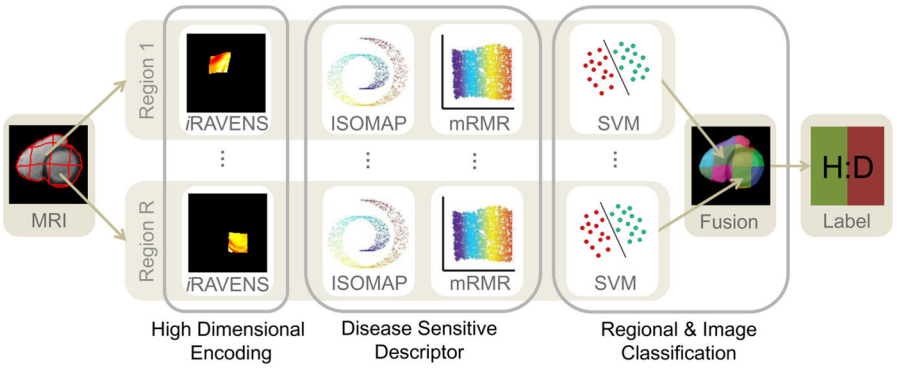
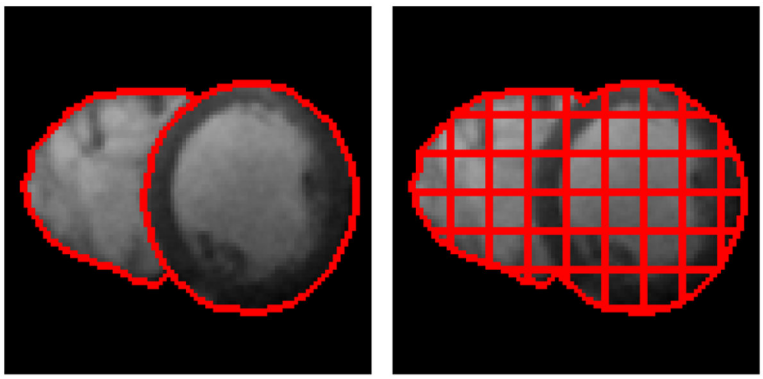


Fig. 1.

An overview of the regional manifold-based classifier. Our approach first computes *i*RAVENS maps (see Section II-A) to encode regional anatomical variations shown in input MRI. From the high-dimensional descriptors, it then infers a low-dimensional encoding of the regional manifolds created by ISOMAP [20] and determines the disease-sensitive embeddings via mRMR [21]. Finally, our approach applies a separate classifier to each regional descriptor and combines the results by weighted fusion to label the input MRI as normal or diseased.



(a) Ventricles (b) Grid

Fig. 2.

Examples of parcellations of the heart into nonoverlapping ROIs. (a) Divides the heart into the two ventricular regions while (b) is based on a regular grid.

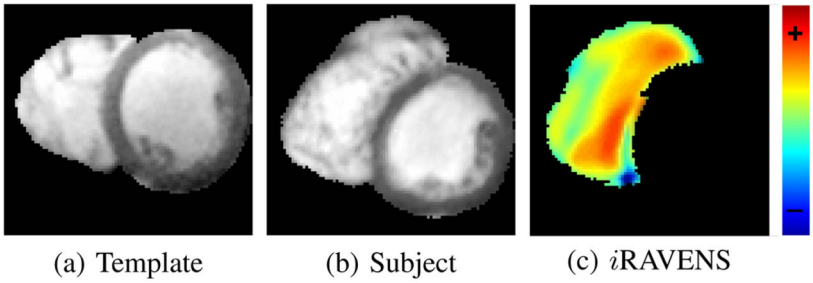


Fig. 3.

Example of *i*RAVENS map: the template image (a) is registered to the scan of a TOF subject (b) resulting in the *i*RAVENS map (c) for the right ventricle region. Red indicates volume expansion while blue indicates shrinkage. Right ventricular dilation of the TOF subject is captured by the volume expansion shown in the *i*RAVENS map in the subject's coordinate space.

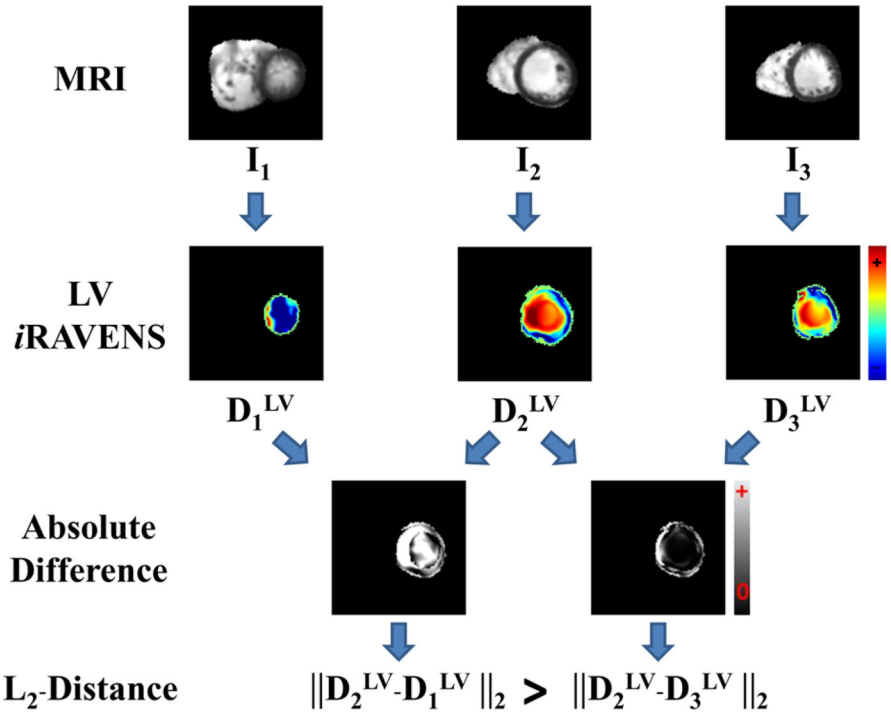


Fig. 4.

Examples of differences between *i*RAVENS maps from different subjects: First row shows the images and second row represents the corresponding *i*RAVENS map of the left ventricle (LV). Differences between *i*RAVENS maps are shown in the third row. For example, I_2 (normal) is less similar to I_1 (TOF) than to I_3 (normal) in the left ventricle, which results in a larger distance $\|D_2^{LV} - D_1^{LV}\|_2 > \|D_2^{LV} - D_3^{LV}\|_2$.

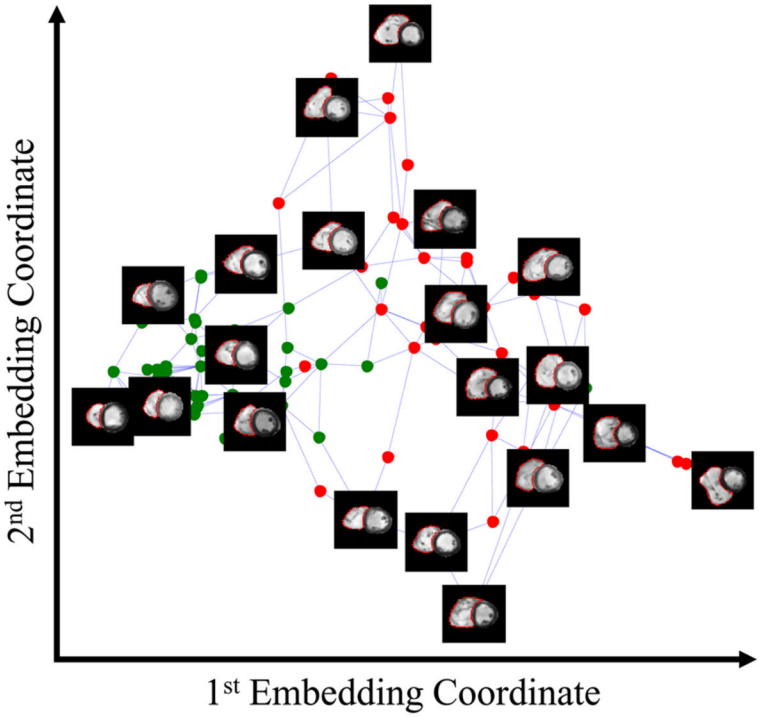


Fig. 5.

Example of 2-D embeddings based on *i*RAVENS in the right ventricular region. Red and green dots represent TOF and normal subjects, respectively. Blue lines connect the nodes with their k nearest neighbors (here: $k = 2$). Only a subset of images is shown to avoid clutter. Right-ventricular boundary of the image is highlighted in red. The neighboring images in the embedding have similar right ventricles in terms of shape and size. In addition, there is a tendency of right-ventricular dilation along the first embedding coordinate from left to right.

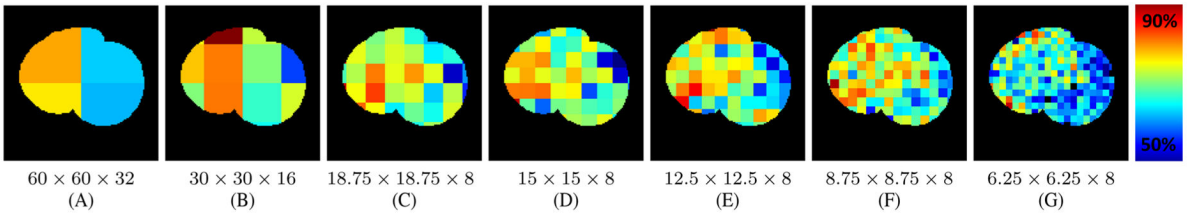
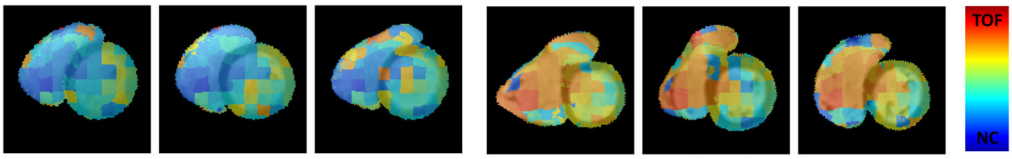


Fig. 6.

Training accuracy of the regional classifiers based on different region sizes: Regions of high weight (red) are around the septum area, which is impacted by reconstructive TOF surgery. Combining the results of the regional classifier based on these weights automatically emphasizes classification results of regions impacted by TOF.



(a) Normal Subjects (b) TOF Patients

Fig. 7.

Examples of abnormality maps for (a) normal subjects and (b) TOF subjects with the grid size being 15 × 15 × 8 mm. Red and blue indicate TOF and normal control (NC) regions, respectively. Maps of normal subjects (a) are generally shown in blue while the maps of TOF subjects (b) mostly indicate the presence of disease with high weights (red).

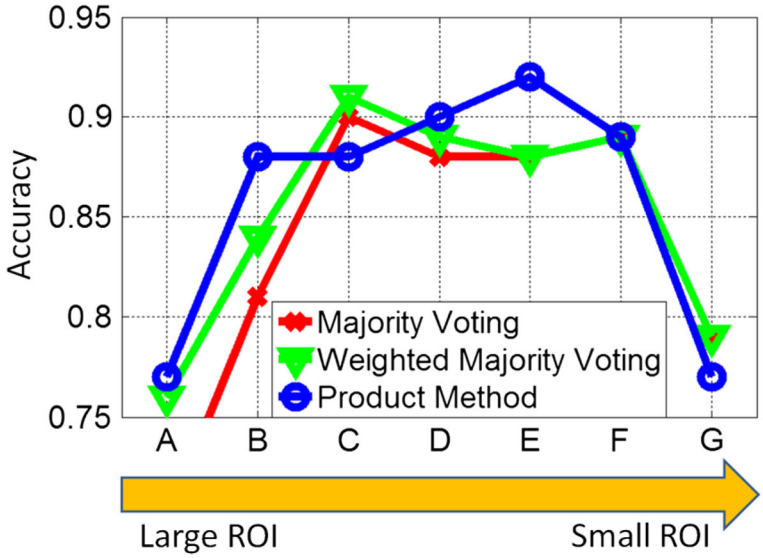


Fig. 8.

Classification accuracy on various size of grid ROIs with different combination strategies (refer Fig. 6 for grid size). There is a significant drop off in the classification accuracy at the largest grid ROI (A) and the smallest grid ROI (G). The product rule (blue circle) outperforms other combination strategies such as majority (red cross) or weighted majority voting (green triangle).

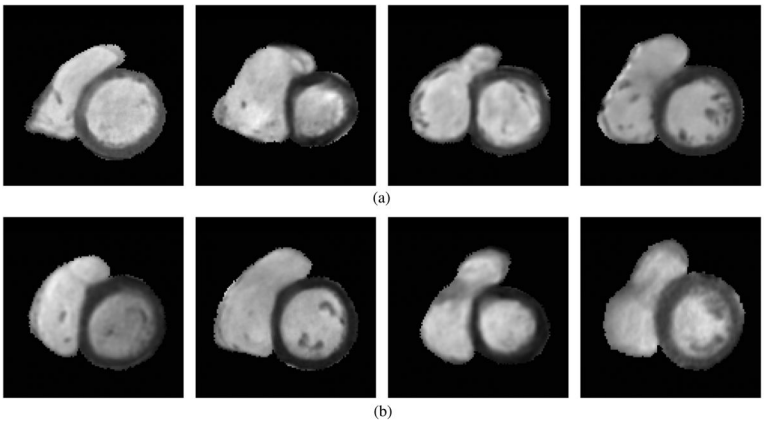


Fig. 9.

Sample slice from each subject that was misclassified by our approach. We note that the misclassified TOF cases are similar to the misclassified healthy subjects.

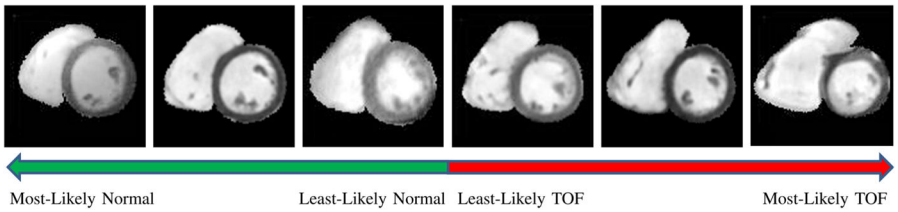


Fig. 10.

Examples of correctly classified scans. Scans were selected based on their overall weight $w_p(1, z)$ as determined by the product rule of (3). The figure illustrates the progression of the shape of the right ventricle from being round to severely dilated and irregular.

Table I

Accuracy of the Disease-Sensitive Encoding Based on the Expert-Defined Partitioning is Much Higher Than the Ventricular Volumes. This Shows That Our Disease-Sensitive Encoding Can Better Describe the Shape Changes Caused by TOF. Encoding Just the Right Ventricle (RV) is Better in Identifying TOF Than Left Ventricle (LV), Which Agrees With the Clinical Understanding of Reconstructive TOF Surgeries Impacting the RV [1], [2]. Also Note That the Accuracy of the Expert-Defined Partitioning Scheme is Lower Then the Grid-Based Scheme, Which was 92%

	Ventricular Volumes	Expert-Defined Partitioning
LV	55%	56%
RV	82%	88%
LV & RV	83%	89%