Ontological labels for automated location of anatomical shape differences

Shane Steinert-Threlkeld, Siamak Ardekani, Jose L.V. Mejino, Landon Todd Detwiler, James F. Brinkley, Michael Halle, Ron Kikinis, Raimond L. Winslow, Michael I. Miller, J. Tilak Ratnamanther

A method for automated location of shape differences in diseased anatomical structures via high resolution biomedical atlases annotated with labels from formal ontologies is described. In particular, a high resolution magnetic resonance image of the myocardium of the human left ventricle was segmented and annotated with structural terms from an extracted subset of the Foundational Model of Anatomy ontology. The atlas was registered to the end systole template of a previous study of left ventricular remodeling in cardiomyopathy using a diffeomorphic registration algorithm. The previous study used thresholding and visual inspection to locate a region of statistical significance which distinguished patients with ischemic cardiomyopathy from those with nonischemic cardiomyopathy. Using semantic technologies and the deformed annotated atlas, this location was more precisely found. Although this study used only a cardiac atlas, it provides a proof-of-concept that ontologically labeled biomedical atlases of any anatomical structure can be used to automate location-based inferences.

1. Introduction

The past several years have seen tremendous advances in cardiovascular imaging [1,2]. Multimodality imaging has permitted characterizing myocardial shape and function in different disease states. Integrating information from these modalities can potentially enhance understanding of cardiovascular morphological and functional response to pathology. As an example, high resolution computed tomography coronary angiography provides detailed information about the vessel anatomy that can be combined with the myocardial perfusion positron emission tomography (PET) or single-photon emission computed tomography (SPECT) scans to evaluate regional blood flow. Integration of data across multiple imaging modalities would benefit from methods that can provide annotations using a consistent nomenclature, as it provides a systematic tool to share information among specialists from different disciplines.

The ability to describe the anatomical structures and their related functions in a formalized manner is critical for sharing data and integrating information across studies. Development of standard anatomical and functional terms and their relationships in a formalized language has been the focus of biomedical ontology. Ontology, traditionally the subfield of philosophy focused with questions about the nature of being, now refers to a mode of formal knowledge representation in the informatics world. In formal ontologies, knowledge about a specific domain can be given to a machine in subject–predicate–object triples so that the machine may perform automated inferences using these statements. Because ontology engineering often requires a well-known and stable body of knowledge to be useful, ontologies have found adoption of varying degrees in a variety of biomedical domains. Ontologies have been designed to describe genes [3], diseases [4], proteins [5], and anatomy [6], among many others. Additionally, an open-access repository of biomedical ontologies called BioPortal [7] has been established by The National Center for Biomedical Ontology [8].

Annotation of medical imaging data is a challenging task as the anatomical information is represented indirectly as image data. Images must be interpreted either manually or automatically to associate image regions with anatomy or physiology. This
segmentation and labeling process can be done per-subject or using a pre-labeled atlas [9,10]. In the atlas-based approach, annotated structures are mapped to a specific study via registration. For instance, in a study of brain activation in schizophrenic subjects (SZs) and healthy volunteers (HVs), ontological labels from the Foundational Model of Anatomy (FMA) were used to annotate neuroimages [9]. Subject images were first coarsely registered to the Talairach atlas [11]. After performing statistical analyses on regions of brain activation, the maximal voxel in each significant cluster was labeled with the appropriate anatomical term from the FMA. These labels were then used to answer questions like, ‘Which parts of the precentral gyrus are active in SZ and HV in these data?’ This and other related work showed how such labels can be used to query the results of neuroimaging studies. Similar approaches can be adapted to annotate imaging data of structures other than the brain in order to perform the same kind of queries.

Despite the healthcare impact of cardiovascular disorders, there have not been many reported studies regarding the annotation of cardiac imaging data. In one study, a diagrammatic representation of congenital heart defects was developed [12]. However, this model is limited to a particular disease state and does not address the integration of information across different studies and imaging modalities. The present study demonstrates a general method for using ontological labels to locate regions of statistical significance in any (including non-brain) anatomical structure. In particular, an annotated atlas of the left ventricle of the human heart was created and combined with the Large Deformation Diffeomorphic Metric Mapping (LDDMM) algorithm [13] to generate diffeomorphic (smooth, invertible, one-to-one, differentiable) transformations which are thought to be superior to other registration techniques [14,15]. Such diffeomorphisms preserve certain topological features like smoothness, connectedness and disjointedness which ensure that all anatomical substructures and many important features are faithfully transferred from one image to the other [16]. For instance, in a segmented image such as an anatomical atlas, all segments are mapped onto the corresponding locations in the target image.

In particular, this study extends the work by Ardekani et al. [17] which used LDDMM and statistical analyses to characterize left ventricular remodeling in the hearts of patients with ischemic cardiomyopathy (ICM) versus nonischemic cardiomyopathy (NICM). The study revealed that nonischemic population had, on average, larger regional myocardial volume relative to the ischemic population. Creating a reference atlas of the left ventricle (LV) with ontological labels provides both a complete automation of the identification of the region of significant tissue volume expansion and a standard frame of reference for future research. More generally, using LDDMM with ontologically labeled biomedical images is shown to provide a fruitful way to automatically locate regions of statistical significance in any anatomical structure.

2. Materials and methods

2.1. Computational cardiac anatomy

Ardekani et al. [17] analyzed in vivo multi-detector computed tomography (MDCT) images of 25 human subjects at both end systole (ES) and end diastole (ED) phases of the cardiac cycle. Of these 25 subjects, 12 had nonischemic cardiomyopathy (NICM) and 13 had ischemic cardiomyopathy (ICM). An average template image at each of ES and ED was generated and then mapped using LDDMM to each target image. Although other results were shown in this previous study, the focus is on the voxel-based Jacobian analysis done using a non-parametric randomized permutation test on all 25 subjects. This Jacobian map encodes local volume difference; a value of 1 indicates none, a value greater than 1 indicates expansion, and a value less than 1 indicates compression. The present study used both a binary version of the ES template and this same template with corrected p values as intensities.

2.2. Ontology extraction

The Foundational Model of Anatomy (FMA) [6] served as our reference ontology. First, the subclass hierarchy of the term Region_of_myocardium was extracted from the full ontology. This extraction was performed using the Jena implementation of sSPARQL [18–20], an extension to the ontology query language SPARQL that allows for recursive and sub-queries. In addition to this extraction, a complete set of terms pertaining to the cardiovascular system in the FMA and other biomedical ontologies has been identified for future extraction.

2.3. Atlas generation

The construction of the imaging atlas for this study began with an ex vivo magnetic resonance image (MRI) of a human heart’s myocardium obtained from the Center for Cardiovascular Bioinformatics and Modeling (CCBM) at Johns Hopkins University (JHU). The imaging protocol was the same as for the canine hearts as detailed in [21].

The image data was converted to ANALYZE7.5 format and hand segmented by a trained clinician according to the 17 parcellation recommendation of the American Heart Association (AHA) [22]. In the AHA parcellation, the left ventricle is divided along the long axis into apical, mid, and basal regions as well as an apex. The apical region is divided along the short axis into four regions, while each of the mid and basal regions is divided into six regions. Fig. 1 shows a schematic of this segmentation.

The hand segmented binary masks were recombined into a single label map image volume, with intensities at each voxel

![Fig. 1. The 17 myocardial parcellation of the left ventricle recommended by the American Heart Association (AHA)](http://www.ccbm.jhu.edu/research/canine_heart_1_060904_normal.php)

---

Please cite this article in press as: Steinert-Threlkeld S et al. Ontological labels for automated location of anatomical shape differences. J Biomed Inform (2012), http://dx.doi.org/10.1016/j.jbi.2012.02.013
corresponding to 10 times the zone number for the particular region of myocardium. The atlas was stored as a NIfTI image with intent code 1002 (i.e. label map). The header field aux_file points to a distinguished text file that maps each intensity to the appropriate term in the extracted ontology. For instance, the first line of the label map is ‘10 http://sig.biostr.washington.edu/fma3.0# Myocardial_zone_1’. The individual segments were recombined and the NIfTI image was built in MATLAB using niftimathlib.3

2.4. Mapping

First, binary images—with intensities 0 and 255—of both the atlas and the ES template were generated. This atlas was first coarsely aligned with the ES template [17] via affine registration using the FSL Linear Image Registration Tool (FLIRT) [23] using 256 bins, 12 degrees of freedom, and trilinear interpolation. Then a four-stage cascading LDDMM mapping [24] was used to diffeomorphically register the linearly deformed binary atlas to the binary ES template. The affine transformation generated by FLIRT was applied to the ontologically labeled atlas and then the LDDMM-generated diffeomorphism was applied to this linearly deformed atlas as illustrated in Fig. 2. This registration also aligned the ontologically labeled atlas with the region of statistically significant volume expansion and the ES template with intensities corresponding to the p values of the analysis of the Jacobian map, both of which were in the same coordinate space as the binary ES template.

2.5. Querying

Once the atlas was mapped onto the average ES template, the ontological labels in the atlas were used to ask four specific questions:

1. What regions of myocardium are annotated in the atlas?
2. What was the average T value of the Jacobian map per region of myocardium?
3. In which region was significant tissue volume expansion observed?
4. What was the average and distribution of the determinant of Jacobian per region of interest?

The third question represents the key to completely automating the diagnosis process in this particular case. Whereas Ardekani et al. [17] used manual expert visual inspection to locate the region of significant volume expansion in the mid anterior, these methods provide the capacity to automatically, formally and precisely ask where the region is located.

These queries were implemented in Java using Jena along with the niftijlib library.4 The source code may be found at http://gitlab.com/shanest/OntologyCV.

3. Theory

The LDDMM algorithm diffeomorphically registers a template image to a target image [13,25]. More precisely, given a background space $\Omega \subset \mathbb{R}^d$, images are defined as functions $I: \Omega \to \mathbb{R}^t$, where for structural images such as MRI, $d = 1$. Given a template image $I_0$ and a target image $I_1$, the goal of image registration is to find a transformation $\varphi: \Omega \to \Omega$ such that $\varphi \cdot I_0 = I_1$.

In the deformable template framework of computational anatomy [26], an exemplar template $I_{\text{empl}}$ is chosen and an anatomy $\mathcal{I}$ is defined as the orbit of $I_{\text{empl}}$ under the diffeomorphism group $\mathcal{G} = \text{Diff}(\Omega)$ action:

$$\mathcal{I} = \{ \varphi \cdot I_{\text{empl}} = I_{\text{empl}} \circ \varphi^{-1} | \varphi \in \mathcal{G} \}$$

Because $\mathcal{I}$ is homogeneous under this group action, for any two $I_0, I_1 \in \mathcal{I}$, there is a $\varphi \in \mathcal{G}$ such that $I_1 = \varphi \cdot I_0$. The goal then is to find the diffeomorphism $\varphi$ connecting two given images of the same anatomical structure.

In the large deformation setting, such a transformation is modeled as the endpoint of a time-indexed flow $\varphi = \phi_t$ associated with a smooth, compactly supported velocity vector field $v: \Omega \times [0,1] \to \mathbb{R}^t$. Writing $v_t$ for $v(\cdot,t)$, the transformation will then be the solution of the ordinary differential equation (ODE)

$$\frac{d}{dt} \phi_t^v(x) = v_t(\phi_t^v(x))$$

(1)

with boundary conditions $\phi_0 = \text{Id}$ and $\phi_1 = \varphi$. Dupuis et al. [27] and Trouvé [28] show that certain smoothness conditions on the vector field ensure that $\varphi$ will be a diffeomorphism.

To compute the diffeomorphism, one first finds the optimal velocity field via the variational problem [13]

$$\dot{v} = \arg \min_{v \in L^2([0,1];\mathbb{R}^t)} \int_0^1 \|v_t\|^2 dt - \frac{1}{\sigma^2} \left\| I_0 \circ \phi_{1,0}^v - I_1 \right\|_{L^2}$$

where $\phi_{1,0}^v = \dot{v} \phi_t^v \circ \phi_0^v$. The resulting Euler–Lagrange equation is

$$2 \nu_t - K \left( \frac{1}{\sigma^2} \left| \nabla \phi_t^v \right|^2 \left( I_t - J_t \right) \right) = 0$$

3 Part of the same project as niftimathlib. See http://niftlib.sourceforge.net/.

Please cite this article in press as: Steinert-Threlkeld S et al. Ontological labels for automated location of anatomical shape differences. J Biomed Inform (2012), http://dx.doi.org/10.1016/j.jbi.2012.02.013
Fig. 3. Because diffeomorphisms preserve submanifolds, the atlas registered to the ES template can be used to locate the region of significant tissue volume expansion. The black region is the statistically significant areas of myocardial tissue volume expansion, superimposed on the LDDMM deformed atlas with ontological labels. The orientation of the heart is similar to Fig. 4.

Table 1  
<table>
<thead>
<tr>
<th>Region</th>
<th>Average T value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myocardial_zone_15</td>
<td>3.40515481</td>
</tr>
<tr>
<td>Myocardial_zone_14</td>
<td>2.8456438</td>
</tr>
<tr>
<td>Myocardial_zone_7</td>
<td>1.7604088</td>
</tr>
<tr>
<td>Myocardial_zone_8</td>
<td>1.4498609</td>
</tr>
</tbody>
</table>

Table 2  
<table>
<thead>
<tr>
<th>Region</th>
<th>Number of voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myocardial_zone_15</td>
<td>696</td>
</tr>
<tr>
<td>Myocardial_zone_10</td>
<td>3</td>
</tr>
<tr>
<td>Myocardial_zone_9</td>
<td>2</td>
</tr>
<tr>
<td>Myocardial_zone_7</td>
<td>1</td>
</tr>
<tr>
<td>Myocardial_zone_4</td>
<td>1</td>
</tr>
</tbody>
</table>

Because diffeomorphisms carry submanifolds to submanifolds and because the 17 segments of the myocardium may be viewed as submanifolds of an LV image, when LDDMM is used to register the ontologically labeled atlas to the ES template, each region of myocardium is carried to the corresponding region on the ES template. Using the ES template with the p values of the analysis of the Jacobian map as its intensities, the labels on the deformed atlas are then used to answer the questions listed in Section 2.5.

4. Results

4.1. Previous results

The voxel-based analysis of the determinant of the Jacobian [17] showed that the NICM group displayed significant myocardial tissue volume expansion in comparison to the ICM group in the mid anterior region. Note that the statistically significant regions were isolated via thresholding and then located (after being superimposed on the ES template) using visual examination only.

4.2. Addressing queries

The answers to the four queries listed in Section 2.5 are given below.

Q1. What regions of myocardium are annotated in this atlas?  
A. Myocardial_zone_1, ..., Myocardial_zone_17.

Q2. What is the average T value for each region of myocardium?  
A. See Table 1.

Q3. Where is the region of statistically significant tissue volume expansion located?  
A. Myocardial_zone_13 = apical anterior. See Table 2 and Fig. 3.

Q4. What was the average and distribution of the determinant of Jacobian per region of interest?  
A. The apical anterior region (Myocardial_zone_13) had the largest difference between the NICM and ICM groups in the average determinant of the Jacobian map as shown in Table 3 and Fig. 4.

5. Discussion

These results show that high resolution biomedical atlases with ontological labels can reliably be used with shape analysis algorithms like LDDMM to automatically locate regions of shape difference in anatomical structures. Although the focus was on remodeling of the left ventricle, this project provides a proof-of-concept because the methods are general and could be adapted to other anatomical structures. One way to extend this work beyond proof-of-concept would be to integrate it into emerging biomedical grid infrastructure such as the Cardiovascular Research
Grid (CVRG)\(^6\) in a manner similar to how the NeuroLex ontology has been developed by and integrated in the Biomedical Informatics Research Network [31].

Although LDDMM was used because of the topological properties of diffeomorphisms mentioned earlier, it has also been shown that high-dimensional image based diffeomorphic mapping methods are the most accurate [15]. While LDDMM was not analyzed in that study, it is regarded as the most advanced such method [14]. The accuracy of the registration used also entails greater precision of the ontological labels [9].

In this study, the AHA recommendation for segmenting the myocardium into 17 regions was used to generate the atlas. While the AHA segments are a well-established standard, they have been mostly visualized by using polar arrangement (as in Fig. 1). Full 3D atlas labeling also provides a method to identify the extent of a pathology within specific segments. More importantly, it could also be used to summarize the measured quantity within the population under study which provides insight into variation of anatomy and function.

The segmentation and labeling used in this study may be most useful for studies that are based on low spatial resolution. More detailed labeling models are required for higher resolution imaging. The advantage of atlas based labeling is that one could potentially define different labeling models using the same atlas and map those to different studies without a need to redesign each specific study. Moreover, because the labeling based on this segmentation is somewhat subjective, in order to use these methods in further studies one would need to generate age and gender specific anatomical and functional cardiac models over different populations. This could be done by using similar labeling techniques and high-dimensional atlases in conjunction with databases such as that being developed by the Cardiac Atlas Project [32].

5.1. Mid versus apical anterior

As mentioned above, Ardekani et al. [17] concluded that the region of interest was in the mid anterior while the ontological methods concluded that this was the apical anterior. These methods also showed that the apical anterior region exhibited the largest difference between the NICM and ICM groups in the average determinant of the Jacobian. This result corroborates the inferred location of the region of interest. The previous study, however, used visual inspection of a trained clinician to locate the ROI. Although it is impossible to say which location is more accurate, the automated approach has the advantage that it can be replicated without expert human interaction. Moreover, the discrepancy between the automated- and expert-driven approaches is small since the two regions border one another and are both in the anterior side of the LV.

5.2. Regional statistics

These methods allow statistical analysis to be done within particular subregions of biomedical images. Through the powerful nonlinear registration of LDDMM, labeled regions from one atlas can be propagated to all subjects and across experiments. Although analyses at the voxel level can be fruitful, for small and subtle pathologies, such analysis may not provide enough power to identify the problem. By looking at statistics at the regional level, one may be able to alleviate this problem.

5.3. Limitations

The current implementation requires high resolution imaging data. There are, however, different imaging modalities such as PET, SPECT, or cardiac MRI that are collected at lower spatial resolution. Each of these modalities contains useful and unique information that needs to be integrated. Although several steps towards

\(^6\) http://cvrgrid.org.

---

Fig. 4. Top panel: Average determinant of the Jacobian map for ICM (left) and NICM (right) populations. Bottom panel: The histogram shows the average of the determinant of the Jacobian in the statistically significant ROI (see Figs. 2 and 3) in each subject. The ICM group (\(n = 13\)) is in blue; the NICM (\(n = 12\)) is in red.

Please cite this article in press as: Steinert-Threlkeld S et al. Ontological labels for automated location of anatomical shape differences. J Biomed Inform (2012), http://dx.doi.org/10.1016/j.jbi.2012.02.013
integrating such information have been made [33], methods are still needed that allow mapping a high resolution atlas to low resolution images from different modalities. As a general tool, LDDMM has the ability to accommodate different modalities. While only high resolution images were used, the methods developed should in principle generalize to other, lower resolution imaging modalities.

Similarly, the methods used in this proof of concept have only been tested on one internal set of data. Although the methods should in theory generalize beyond the data used, it is for this reason that this paper represents a proof-of-concept. As mentioned earlier, integrating the tools developed with emerging biomedical computing grids would enable verifying these methods with data from other studies and other imaging modalities.

For future applications, integration with neuroimaging software such as Slicer3D [34,35] will be necessary. For example, clinical data could be mapped to the annotated atlas to facilitate comparison of patient-specific data with the population statistics in the different regions. Further annotation of disease specific atlases is possible. Such integration will provide a friendly querying interface and is a current focus of research and development.

Acknowledgments

This research was funded by a Provost Undergraduate Research Award (Office of the Provost, Johns Hopkins University), NIH R24 HL085343, NIH P41 EB015909-12, and NIH HL087706. We would like to thank Geoffrey Gunther for assistance on the segmentation library and related technical assistance.

This research was funded by a Provost Undergraduate Research Award (Office of the Provost, Johns Hopkins University), NIH R24 HL085343, NIH P41 EB015909-12, and NIH HL087706. We would like to thank Geoffrey Gunther for assistance on the segmentation library and related technical assistance.

Please cite this article in press as: Steinert-Threlkeld S et al. Ontological labels for automated location of anatomical shape differences. J Biomed Inform (2012), http://dx.doi.org/10.1016/j.jbi.2012.02.013

References


