Automatic classification framework for ventricular septal defects: a pilot study on high-throughput mouse embryo cardiac phenotyping

Zhongliu Xie
Xi Liang
Liucheng Guo
Asanobu Kitamoto
Masaru Tamura
Toshihiko Shiroishi
Duncan Gillies

Automatic classification framework for ventricular septal defects: a pilot study on high-throughput mouse embryo cardiac phenotyping

Zhongliu Xie,^{a,b,*} Xi Liang,^{b,c} Liucheng Guo,^d Asanobu Kitamoto,^b Masaru Tamura,^{e,f} Toshihiko Shiroishi,^e and Duncan Gillies^a

^aImperial College London, Department of Computing, South Kensington Campus, London SW7 2AZ, United Kingdom

^bNational Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

^cUniversity of Melbourne, Department of Computer Science and Software Engineering, Parkville Campus, Melbourne VIC 3010, Australia

^dImperial College London, Department of Electrical and Electronic Engineering, South Kensington Campus, London SW7 2AZ, United Kingdom

^eNational Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan

^fRIKEN BioResource Center, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Abstract. Intensive international efforts are underway toward phenotyping the entire mouse genome by modifying all its $\sim 25,000$ genes one-by-one for comparative studies. A workload of this scale has triggered numerous studies harnessing image informatics for the identification of morphological defects. However, existing work in this line primarily rests on abnormality detection via structural volumetrics between wild-type and gene-modified mice, which generally fails when the pathology involves no severe volume changes, such as ventricular septal defects (VSDs) in the heart. Furthermore, in embryo cardiac phenotyping, the lack of relevant work in embryonic heart segmentation, the limited availability of public atlases, and the general requirement of manual labor for the actual phenotype classification after abnormality detection, along with other limitations, have collectively restricted existing practices from meeting the high-throughput demands. This study proposes, to the best of our knowledge, the first fully automatic VSD classification framework in mouse embryo imaging. Our approach leverages a combination of atlas-based segmentation and snake evolution techniques to derive the segmentation of heart ventricles, where VSD classification is achieved by checking whether the left and right ventricles border or overlap with each other. A pilot study has validated our approach at a proof-of-concept level and achieved a classification accuracy of 100% through a series of empirical experiments on a database of 15 images. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.2.4.041003](https://doi.org/10.1117/1.JMI.2.4.041003)]

Keywords: mouse embryo phenotyping; ventricular septal defects; atlas-based segmentation; snake evolution.

Paper 15073SSPRR received Apr. 1, 2015; accepted for publication Jul. 30, 2015; published online Sep. 11, 2015.

1 Introduction

For many decades, scientists have been studying genetic impacts on mammalian development and disease pathology. The Human Genome Project,¹ for example, was a grand milestone marking the completion of human DNA sequencing and gene mapping, since which researchers have carried on to study gene-associated functional information. This process, known as phenotyping as compared to genotyping, however, is very difficult to follow due to numerous issues regarding modifying human genes, especially the moral concerns. The mouse genome, therefore, has widely been used as a surrogate model, due, in part, to the findings that 99% of mouse genes have a homologue in the human genome,² as well as other advantages including the low cost, short time required for breeding, etc. Intensive global efforts, such as the International Knockout Mouse Consortium³ and International Mouse Phenotyping Consortium (IMPC),^{4,5} have been underway toward understanding all the $\sim 25,000$ mouse genes by systematically modifying them one-by-one for comparative analysis. More specifically, an entire series of mouse embryonic stem cells is under development with each gene to be mutated using, e.g., gene targeting techniques⁶ to generate a tailored mouse strain, which is then raised in order to study

the impact of gene-mutation with respect to morphology, metabolism, or other biological traits (also known as phenotype). However, with vast amounts of data to analyze, the traditional means via time-consuming histological examination faces significant challenges,⁷ which in turn have stimulated a large-scale exploration for high-throughput approaches. Biomedical imaging technology, such as CT and MRI, thus has started being used to facilitate phenotyping practices concerning morphological abnormality, especially with the employment of image informatics to automate defect recognition.

In terms of morphological phenotyping, existing image informatics work primarily rests on volumetric comparison of anatomical structures between wild-type and gene-modified mice for abnormality detection.^{8–12} In these approaches, an atlas is typically created in advance with target structures manually labeled out by experts. At test time, target images are registered to the atlas for backpropagation of labels to perform segmentation, which is widely known as atlas-based segmentation or segmentation propagation. In practice, state-of-the-art performance normally involves the use of multiple atlases together for segmentation, which is often referred to as multitlas label propagation, and there is an abundance of such algorithms in

*Address all correspondence to: Zhongliu Xie, E-mail: zx10@imperial.ac.uk

the literature.¹³ Once structures of interest are segmented, volumetric evaluations can then be performed to separate abnormal subjects from the normal ones. However, volume contrast merely achieves a superficial level of screening for morphological abnormality and generally fails when the pathology involves no severe volume changes. Another important challenge facing the segmentation propagation framework is the access to a suitable atlas. Despite the considerable progress on its applications to the adult mouse model, especially segmentation of the mouse brain^{14–20} and the relatively easy access to adult mouse brain atlases,^{11,21–23} embryo phenotyping, in contrast, has encountered significant obstacles. This is largely because mouse embryo development typically lasts for only ≈ 18.5 days with one day difference potentially leading to dramatic changes due to rapid organogenesis. In addition, nurturing conditions can also make a difference, which collectively makes it very difficult to use public atlases in practice. On the other hand, the availability of public atlases is actually very limited, with the only ones, to the best of our knowledge, being an E15.5+ MRI atlas⁸ and an E15.5 μ -CT atlas,²⁴ both of which are incompatible, for example, to the E14.5 embryos used in this work, let alone the issues concerning imaging protocol. In spite of these difficulties, it is estimated that $\sim 30\%$ mutant mouse lines would end up as embryonic lethal,²⁵ which has drawn considerable attention to prenatal phenotyping.

In the domain of prenatal phenotyping, congenital heart diseases (CHDs) have triggered wide-spread concerns, yet there is very limited research progress harnessing image computing technology to the classification of heart phenotypes. This is not only because of the lack of embryonic heart segmentation work (where only a few studies on semiautomatic segmentation are found²⁶), but heart defects also tend to be more subtle and often difficult to classify via volumetric contrast. The ventricular septal defect (VSD), for instance, is one of the most common CHDs among others including cardiomyopathy and atrial septal defects.²⁷ This particular phenotype is featured by the presence of a hole or similar defect in the cardiac ventricular septum that divides the left and right ventricles, for which no severe heart volume difference may be observed across the healthy and defective populations. Figure 1 illustrates the pathology with a comparison between VSD positive and negative subjects under μ -CT scanning. To address problems of this type, another school of researchers have suggested employing morphometrics based on a range of deformation features emerging during image warping to the atlas, in order to detect anomalous

structures/areas without the need of heart segmentation. Notable deformation features include positional shift,¹² Jacobian determinant,^{12,28} deformation stress,²⁸ etc. However, such detection is not robust since deformation features could vary significantly under different registration settings and using different reference templates. Moreover, a common limitation of both volumetric and deformation-based approaches, owing to the complex nature of phenotypic characteristics, is that instead of directly classifying phenotypes, they merely serve the purpose of abnormality detection, where the suspicious structures/regions detected then generally require a follow-up manual inspection by domain experts to identify the actual phenotype. For that reason, current image-based practices on embryo cardiac phenotyping are still limited to the role of facilitating manual assessment with enhanced visualization of cardiovascular structures and malformations via, e.g., the optimization of imaging protocol,²⁹ the use of a tissue staining technique,³⁰ three-dimensional (3-D) image reconstruction for cavity analysis,³¹ metric analytics based on the diameter of great arteries and semilunar valves,³² etc. The involvement of expert manual labor has been a critical bottleneck that restricts such practices from meeting the high-throughput requirement, leading to the demand for more sophisticated phenotyping frameworks.

This paper presents an extended version of our earlier prototype,³³ which, to the best of our knowledge, is the first fully automatic VSD classification system in mouse embryo image computing. The extension provides a complete methodology starting from a data acquisition protocol that follows the international standards, assuring its wide applicability in this community. Then in terms of image processing, the derived μ -CT raw images first go through a mouse embryo extraction process to remove the embryo from its surroundings, followed by the creation of a local atlas using all the wild-type subject images. At test time, a rigid-affine-nonlinear three-step registration scheme is applied to align the embryo in a target image with the atlas in order to perform a standard atlas-based segmentation on the heart and then on the ventricles; the ventricle segmentations are then further refined using a snake evolution (also known as active contours) technique. VSD classification is completed by checking whether the left and right ventricle segmentations border or overlap with each other. The key novelty of our approach, from a methodological perspective, is the combined use of atlas-based segmentation and snake evolution, which is able to segment an unseen target image in the presence of pathology and attain a high VSD classification accuracy with the use

Fig. 1 Ventricular septal defects (VSD) pathology: (a) a negative/healthy subject image and (b) a positive/defective sample. The left and right ventricles are separated in the former case, but are interconnected in the latter case.

of only one atlas throughout the procedure. The minimal atlas requirement in turn provides a significant advantage that lowers the cost of local atlas creation to a minimum level in case of an inability to obtain a public atlas. Furthermore, a new snake evolution algorithm has been developed in this work which has managed to improve VSD classification accuracy to 100% from the 90.9% scored in the prototype system³³ on a pilot study using a database of 15 images.

2 Methods

2.1 Subject Preparation and Imaging Protocol

All of our animal experiments are approved by the Animal Care and Use Committee of National Institute of Genetics (NIG, Mishima, Japan). The mouse strains used in this study are maintained at the Genetic Strains Research Center, NIG, in a specific pathogen free facility (12 h light and dark cycles). Embryos are obtained from natural mating. Noon on the day when the vaginal plug is detected is designated as the 0.5 day post coitum (dpc) point. All mouse embryos in this work are collected at 14.5 dpc, washed in phosphate buffered saline solution, and maintained in 4% paraformaldehyde solution until μ -CT scanning. Right before scanning, the embryos are soaked in a contrast agent made by a 1:3 mixture of Lugol solution and double distilled water in advance. These procedures are carefully carried out to make sure no air flows into the heart ventricles and blood in the ventricles is flushed as much as possible. All the samples are then scanned, in our case, using a SCANXMATE-E090S 3-D μ -CT machine (Comscan Techno, Japan), during which each embryo is fitted in a 1.5 ml Eppendorf tube and fixed by wet paper. The x-ray radiation is applied at a tube voltage peak of 60 kVp and a tube current of 130 μ A. The imaging subject is rotated by 360 deg at 0.36 deg per step, generating 1000 projections at a 640×480 pixel resolution, and reconstructed in 3-D at an isotropic resolution of $25 \times 25 \times 25 \mu\text{m}^3$. Figure 2 shows an example raw image and its intensity histogram, with the mouse embryo contained in a test tube (circle-shaped object in axial view) and soaked in the maintenance solution (slightly brighter material between mouse and tube wall).

2.2 Mouse Embryo Extraction

The first major computational process in the pipeline is to extract embryo from its surrounding liquid solution, glass tube, and the air inside/outside the tube. A simple histogram

study reveals that signals tend to be distributed into several clusters, with the first three being air, tube, and solution, respectively (denoted k_1, k_2, k_3), while the residuals that collectively compose the mouse (k_4), have a bizarre distribution pattern due to different tissue reflection properties. Given this, we perform the following:

Multiscale histogram smoothing: Denote the number of bins as N_b ; the histogram is smoothed at multiple scales with a radius factor $a^{(i)}$ at each stage i , and the voxel count $\theta_b^{(i)}$ of bin b each time is updated by averaging the values spanning over the interval $[b - a^{(i)} \cdot N_b, b + a^{(i)} \cdot N_b]$. The histogram would then become much less steep, with most of the local maxima and minima smoothed out.

Cluster localization and thresholding: Next, we use first- and second-order derivatives to localize the clusters by drawing the first four peak positions, respectively P_1, P_2, P_3, P_4 , from the set $S_p = \{b | (d\theta_b/db) = 0, (d^2\theta_b/db^2) < 0\}$ in ascending order, followed by a binary thresholding to separate k_4 from k_1, k_2, k_3 by setting a threshold to the valley between P_3 and P_4 , or the first valley following P_3 .

$$\Phi_1 = \min \left\{ b \mid \frac{d\theta_b}{db} = 0, \frac{d^2\theta_b}{db^2} > 0, b > P_3 \right\}. \quad (1)$$

In some rare occasions, as is the case of two datasets in our sample pack, clusters k_3 and k_4 may be closely located and the signals interfere with each other, causing the supposed P_4 to be overwhelmed in k_3 , failing in Eq. (1) to derive the desirable threshold. A closer observation in this case will discover k_3 to be steeper than k_4 , implying there is a point at which gradient descent suddenly slows down as it transits from k_3 to k_4 , and it can be captured via second- and third-order derivatives.

$$\Phi_2 = \min \left\{ b \mid \frac{d^2\theta_b}{db^2} = 0, \frac{d^3\theta_b}{db^3} < 0, b > P_3 \right\}. \quad (2)$$

A combined approach without knowing the case information is simply $\Phi = \min\{\Phi_1, \Phi_2\}$.

Embryo extraction: After thresholding, the largest foreground region is extracted, while smaller regions, mostly being k_3 voxels (with an intensity above Φ) misidentified as k_4 , are relabeled as background. The label map is then dilated, e.g., with a ball-shaped kernel, to reinclude k_4 voxels (with intensity below Φ) that are accidentally ruled out during thresholding. The final label map is applied to the image, with background voxels reassigned to $\gamma = P_3$ and the foreground untouched.

Fig. 2 Example raw image of mouse embryo (a) in axial view, (b) in sagittal view, and (c) its intensity histogram.

2.3 Atlas Creation Using Wild-Type Mean Image

The label atlas creation job rests on the construction of a mean image using all the wild-type control images available at hand and follows the procedure sketched in Fig. 3. The purpose of using wild-type subjects only in this case is to reduce the effect of variations that are not caused by gene removal. As the first step, one of the images is randomly selected as the initial reference and the rest are registered and warped toward it using a rigid registration scheme. Subsequently, all the rigidly aligned images are then averaged to generate a new reference, which is blurry but unbiased toward the geometry of the initial reference. The aligned images are then warped to the rigid-mean reference via affine registration, followed by further averaging to update the reference. After these two steps, a global correspondence regarding embryo positions, scales, and orientations across the control image pool should have been secured. Then a B-spline nonlinear registration is iteratively applied to locally align them with the reference, which is constantly updated after each run via a similar averaging process, until boundaries become sharp and anatomical structures become clear. The implementation of registration schemes in our case is primarily based on the open source Image Registration Toolkit repository,³⁴ the energy function used in the nonlinear registration leverages Mattes mutual information,³⁵ and a rigidity penalty.³⁶ Furthermore, a multiresolution scheme is also used with a down-sampling factor of 0.5 at each dimension, and the spacing of control points gradually reduces from 1.6 to 0.8, 0.4, and 0.2 mm. Once the final mean image is generated, the heart as a whole and its internal left and right ventricles are manually annotated by a domain expert to create the label atlas.

2.4 Heart Segmentation

With the label atlas at hand, heart segmentation is then performed on the target embryo images. This process follows the standard routine of atlas-based segmentation widely covered in the literature. Simply put, a three-step alignment scheme consisting of consecutive rigid, affine, and nonlinear registrations is employed to warp a target image to the atlas space, with the same parameter settings as above. Once the target image is non-linearly aligned with the atlas, labels of the heart are then back-propagated to the target image through reverse transformation. Some postprocessing, such as label map erosion and dilation, may also be carried out for morphological correction purposes, depending on the quality of the segmentation outcome. A notable point is that only a relatively coarse segmentation accuracy is required at this stage, as the heart segmentation merely

serves as an initialization mask for the ventricle segmentation that comes next.

2.5 Ventricle Segmentation

To segment heart ventricles, we start with another atlas-based segmentation by applying a similar nonlinear registration to align the target image with the atlas again, this time with a mask over the heart region. The mask is derived from the previous heart segmentation with a padding of an additional 5 voxels in each direction. The labels of the ventricles in this case are backpropagated to generate an initial segmentation of the left and right ventricles, respectively. However, segmentation accuracy at this point would be limited, due to the difficulty of perfectly registering a morphologically defective heart to a healthy atlas, let alone the nature of inter-subject variability. Therefore, a further refinement using snake evolution follows, as described below.

The snake (also known as active contour) model is a classic segmentation technique originally proposed by Kass et al.³⁷ It works by actively evolving the contour overtime under a collection of forces. With borrowed insights from the Region Competition approach,^{38,39} in our work, we project the image into a feature space F using a Gaussian mixture model (GMM) to form the internal deformation field, whereas the classic curvature model³⁷ C is adopted to model the external force. Our implementation of a snake evolution framework builds on top of the open source ITK-Snap repository,^{39,40} while the actual algorithm is based on the following theory. Denoting the arcs composing the snake/contour as s , the energy function is formulated as

$$E = \int \{[\alpha \cdot F(s) - \beta \cdot C(s)] \cdot \vec{p}_s\} ds, \quad (3)$$

where $F(s)$ and $C(s)$ are, respectively, the internal/feature and external/curvature forces, weighted correspondingly by parameters α and β , and $\vec{p}_s$ is the normal vector perpendicular to s pointing outward. In this particular problem domain, heart ventricles are filled with blood when the mouse embryo is alive. After death, since the embryo is washed and maintained in the liquid solution described earlier, blood will mostly be replaced by the solution, except for some residuals that still remain. As a result, the ventricles are jointly filled with both fluids, which are mutually insoluble. Under μ -CT scanning, blood appears brighter and solution darker, collectively surrounded by cardiac muscles of some intermediate brightness, all at a relatively consistent intensity level across different

Fig. 3 Creation of the wild-type mean image through a sequence of group-wise image alignment and averaging processes with rigid, affine, and nonlinear registrations using all the available wild-type images.

images. In this case, ventricle segmentation is delegated to the labeling of such mixed fluids, and we formulate the feature space as

$$F(s) = \max[G_{\mu_1, \sigma_1}(s), G_{\mu_2, \sigma_2}(s)] - \psi, \quad (4)$$

where each of the two Gaussians, G_{μ_1, σ_1} and G_{μ_2, σ_2} , captures one type of fluid, as illustrated in Fig. 4. A positive $F(s)$ leads to a local expansion and is negative for contraction. Parameters $\langle \mu_1, \sigma_1, \mu_2, \sigma_2 \rangle$ should be empirically determined according to the imaging protocol and reflective property of the maintenance solution used, and ψ is the expansion-contraction threshold normally set to 0.5.

The key to a successful snake evolution practice is snake initialization, where we return to the coarse ventricle segmentations obtained by the atlas-based label propagation earlier. The initialization seed is derived from a minor level of erosion on the segmentation and snake evolution is applied separately to each ventricle separately. The evolution runs over a number of iterations, during which the snake keeps evolving based on collective forces, until an equilibrium is reached or a maximal number of iterations carried out. In many occasions, an initial snake from an (inaccurate) atlas-based segmentation may extend to surrounding heart muscles and/or not cover the whole ventricle. During evolution, the contour inside the ventricle will keep expanding while its counterpart in the surrounding myocardial regions will keep contracting, until the termination condition is met.

Fig. 4 Feature function superimposed with the intensity histogram of heart region: each voxel is projected to the feature space and derives a value based on its intensity with respect to the feature function.

2.6 VSD Classification

Once snake evolution is complete on both ventricles, VSD classification is then carried out by checking whether the two snakes (final segmentations) border or overlap with each other. If so, ventricular connectivity is concluded (i.e., a hole is present) and the subject is classified as VSD positive. Otherwise, the subject is classified VSD negative.

3 Results and Discussions

3.1 Test Data

The data acquisition process was carried out by an expert team composed of biologists from NIG and RIKEN BioResource Center. There were 15 mouse embryos collected in this study, consisting of 3 healthy wild-type C57BL/10 samples and 12 mutant subjects through recombination-induced mutation 4 (Rim4).⁴¹ Among the mutants, three were homozygotes, which were generated by inbreeding the Rim4 mouse lines, whereas the remaining nine were heterozygotes, generated by cross-breeding between the wild-type and Rim4 populations. An expert manual assessment via histological examination was carried out, and all three homozygous samples were confirmed to be VSD positive, while all the heterozygous and wild-type subjects were VSD negative. In addition, a simple Student's *t* test was applied, and it was concluded that there is no statistically significant difference ($p > 0.05$) of the heart volume measures between every two genotype groups and between VSD positive/negative samples, demonstrating a clear infeasibility with volumetric approaches.

3.2 Embryo Extraction Results

The embryo extraction results were assessed qualitatively by domain experts and found to be satisfactory, in particular, a 3-D model was reconstructed for each extracted mouse image to facilitate evaluation. An example of a raw image of a mouse embryo superimposed with a label map generated by the algorithm is shown in Fig. 5(a); this was used to obtain the extracted embryo in Fig. 5(b), which was then successfully portrayed in 3-D with a high level of detail, including head, body, and tail as well as the number of fingers and toes, as shown in Fig. 5(c). Quantitative measures (such as the Dice score), however, were difficult to compute since there was no

Fig. 5 Example mouse embryo extraction result: (a) a raw image superimposed with the label map by our algorithm, (b) the embryo image after extraction, and (c) three-dimensional reconstruction of the extracted embryo.

ground truth information available. On the other hand, such measures are not part of the goal here for VSD classification.

In our best experiment performance, we set $N_b = 1000$ and chose a three-stage processing with $a^{(i)} = 1/25, 1/35$, and $1/45$, respectively, for the multiscale histogram smoothing. This, however, is a simple guideline since there could be numerous settings to achieve this task, as long as most local maxima/minima are smoothed out and the signal distribution of the clusters follows a Gaussian-like pattern. In fact, a more straightforward idea for cluster localization may be to fit a GMM model as used in the human brain tissue classification.⁴² The GMM is expected to capture the mean μ_k and standard deviation σ_k of each cluster; once obtained, the threshold can then be simply set to $\Phi = \mu_3 + 2\sigma_3$ or similar values. However, this does not work well in practice, as the mouse cluster does not follow a simple Gaussian distribution. Modeling a mouse as another Gaussian mixture is also difficult, as there is no way of knowing how many subclusters would fit the data.

3.3 Atlas Creation Results

The interim result right after group-wise linear processing was very blurry, yet it was quickly refined as nonlinear alignment took place. After five iterations of nonlinear processing, boundaries have already become sharp, and after eight/nine iterations, improvement becomes marginal. The reference templates right after linear processing, and after 5, 8, and 10 iterations of nonlinear processing are shown in Fig. 6. A notable point is that abdominal structures will always be more blurry due to substantial natural variations. In our work, we carried out 10 iterations before finalizing the template, which was then submitted to domain experts for label annotation. The atlas created in this study is shown in Fig. 7, including both a heart label map and a ventricle label map.

In fact, warping target images to a representative template space for comparative analysis is by all means not new; rather it is a common practice in biomedical image computing. There have been a number of paradigms proposed for unbiased atlas creation,^{24,28,43} more or less akin to ours. The key insight of its inclusion here is to adapt such a method to the field of embryonic VSD classification as well as to secure the completeness of our methodology for a perfectly reproducible result. In addition, given the difficulty to secure a suitable public atlas as described earlier, an ideal embryo phenotyping practice often entails the use of local atlases. A minimal atlas requirement in this case is

Fig. 7 (a) A two-dimensional (2-D) image slice of the created heart atlas: the red-colored area represents the heart. (b) A 2-D slice of the created ventricle atlas: the green and red areas, respectively, represent the left and right ventricles.

highly desirable as atlas creation can be both expensive and time-consuming. Since our methodology utilizes a combination of atlas-based segmentation and snake evolution to achieve VSD classification, only one atlas is required, which is significantly superior to approaches using multiple atlases.

3.4 Heart Segmentation Results

Due to the use of only one atlas for image registration and label backpropagation, the heart segmentation was restricted to a relatively coarse accuracy, especially on boundary regions. Figure 8 shows some example segmentation outcomes on four target images, two of which are VSD positive while the rest are VSD negative samples. As described earlier, heart segmentation (after fitting with a small padding) merely serves as an initialization mask for ventricle segmentation; thus, its accuracy is not critical in this case. On the other hand, such a prestep is essential, as a direct ventricle segmentation on defective subject images would lead to lower accuracy, and a snake evolution algorithm confined to this region can only achieve a higher efficiency and prevent accidental expansion to areas outside the heart.

3.5 Ventricle Segmentation and Ventricular Septal Defect Classification Results

As described earlier, in this stage, an atlas-based label propagation was applied to all 15 test datasets first, generating a preliminary ventricle segmentation for each image, which after minor erosion, were then used as the initialization seed for snake

Fig. 6 Template updates (a) right after group-wise linear alignment and averaging, (b) after 5 iterations, (c) 8 iterations, and (d) 10 iterations of nonlinear processing.

Fig. 8 Example heart segmentation results of four target images: (a) and (b) two VSD positive samples, and (c) and (d) two VSD negative samples. N.B. only a relatively coarse accuracy is attained in this stage, as heart segmentation merely serves an initialization mask for ventricle segmentation.

evolution. Snake evolution was tested with the following six different settings; each was carried out twice to confirm its reproducibility.

1. $\langle \alpha = 0.9, \beta = 0.1, \max Itr = 500 \rangle$
2. $\langle \alpha = 0.9, \beta = 0.05, \max Itr = 500 \rangle$
3. $\langle \alpha = 0.9, \beta = 0.2, \max Itr = 500 \rangle$

$$4. \langle \alpha = 0.6, \beta = 0.1, \max Itr = 500 \rangle$$

$$5. \langle \alpha = 0.6, \beta = 0.05, \max Itr = 500 \rangle$$

$$6. \langle \alpha = 0.6, \beta = 0.2, \max Itr = 500 \rangle$$

In general, the algorithm demonstrated a considerably high robustness with the same VSD classification outcome almost all the time, despite some small discrepancies in the generated label maps. Figure 9 illustrates an example outcome with setting 1 for three cases, respectively, a VSD positive and two negative cases. A notable point is that blood residuals left in the ventricles during data acquisition can make a major impact, as shown in Figs. 9(b) and 9(c). In the former case, the residuals were mostly flushed away, which enables the contour/snake to readily expand from one ventricle to the other, resulting in almost entirely overlapping segmentations. In the latter case, there was a significant amount of blood residual, which produced a higher barrier for expansion, leading to a low level of overlap after snake evolution. Moreover, a rare false negative was generated for this particular sample with setting 6, due to a comparatively stronger curvature force that restricts snake expansion, as shown in Fig. 10, which is similar to the situation that occurred with our earlier prototype.³³ To sum up and in statistical terms, our algorithm successfully achieved a 100% accuracy in settings 1 to 5 and a 93.3% accuracy in setting 6. However, given only the small volume of data used in this pilot study, the true accuracy remains uncertain and requires an extensive investigation on larger databases.

3.6 Runtime Environment and Processing Time

The runtime environment was deployed on a standard PC with an Intel i7 3.4 GHz quad-core CPU and 16 GB RAM memory without parallel settings. Including all runs of experiments on the whole test database collectively, mouse embryo extraction

Fig. 9 Ventricle segmentation and VSD classification results: (a) and (c) a VSD negative and two positive cases, and (d) to (f) the images superimposed with corresponding label maps. Fluid in left/right ventricle is colored in red/blue accordingly. N.B. in (b), the blood was mostly flushed away during data acquisition, which enables contour/snake to readily expand from one ventricle to the other, resulting in almost entirely overlapping segmentations, whereas in (c), there was a significant amount of blood residuals, leading to a low level of contour/snake overlap after evolution.

Fig. 10 Results of snake evolution on image (a) with the settings (b) $\alpha = 0.9, \beta = 0.1, \max ltr = 500$ and (c) $\alpha = 0.6, \beta = 0.2, \max ltr = 500$. The former successfully identified ventricular connectivity, while the latter did not.

on average took 42 s, atlas-based segmentation on the heart and then ventricles (with mask over the heart region), respectively, 11 and 9 min, followed by 30 s of snake evolution, i.e., ~21 min of computation time per image, which is considered a substantially high throughput. The efficiency could potentially be even higher using parallel computing. On the other hand, atlas creation may take several hours to complete, depending on data size and number of iterations used during creation. Fortunately, this can be done in advance and once for all applications.

4 Conclusion and Future Work

In this study, we proposed, to the best of our knowledge, the first fully automatic classification framework for VSDs in mouse embryo cardiac phenotyping (except for the atlas labeling part, which requires expert manual annotation but is done only one time). The data preparation and imaging protocol used in our work conforms to the prevailing IMPC standards; therefore, our methodology is widely applicable in this field. The complete image computing algorithm starts with the extraction of the mouse embryo from its surroundings in the raw image, where the wild-type datasets are then used to construct an unbiased mean image for the creation of a representative label atlas. Subsequently, an atlas-based label propagation follows to perform segmentation on the target image over the heart and then ventricle regions. The derived ventricle segmentation then separately triggers a snake evolution algorithm in each ventricle. As the last step, VSD classification is carried out by checking whether the two contours/snakes border or overlap with each other. A series of experiments have been run on a database of 15 mouse embryo images, three out of which were VSD positive and the rest were VSD negative samples. Our algorithm successfully achieved an overall accuracy of 100% on five out of six experimental settings, with only one false negative classification generated on the remaining setting, which lowered the accuracy to 93.3% in this particular case.

A major limitation with this pilot study is the relatively small size of the test database for experimentation; hence, this study has only validated our VSD classification framework at a proof-of-concept level. The true accuracy remains uncertain when applied to larger data volumes, calling for further investigations. Moreover, an important insight regarding our methodology drawn from this study is the impact of an optimized data acquisition process on the follow-up image processing operations, in the sense that one should try to eliminate air and blood residuals in the ventricles as much as possible before the imaging procedure. Further studies should be carried out to explore the optimization of the data preparation protocol in more depth. In

addition, although this study was targeted at VSD classification, other phenotypes/defects that are similar in nature (e.g., atrial-septal defects) may be addressed in the future using extensions of our method.

References

1. National Human Genome Research Institute, "An overview of the human genome project," 2012, <http://www.genome.gov/12011238> (19 August 2015).
2. Mouse Genome Sequencing Consortium et al., "Initial sequencing and comparative analysis of the mouse genome," *Nature* **420**(6915), 520–562 (2002).
3. F. S. Collins et al., "A mouse for all reasons," *Cell* **128**(1), 9–13 (2007).
4. International Mouse Phenotyping Consortium (IMPC), www.mousephenotype.org
5. S. D. M. Brown and M. W. Moore, "Towards an encyclopaedia of mammalian gene function: the International Mouse Phenotyping Consortium," *Dis. Models Mech.* **5**(3), 289–292 (2012).
6. M. R. Capecchi, "Generating mice with targeted mutations," *Nat. Med.* **7**(10), 1086–1090 (2001).
7. S. D. M. Brown, J. M. Hancock, and H. Gates, "Understanding mammalian genetic systems: the challenge of phenotyping in the mouse," *PLoS Genet.* **2**(8), 1131–1137 (2006).
8. J. O. Cleary et al., "Magnetic resonance virtual histology for embryos: 3D atlases for automated high-throughput phenotyping," *NeuroImage* **54**(2), 769–778 (2011).
9. F. C. Norris et al., "Segmentation propagation using a 3D embryo atlas for high-throughput MRI phenotyping: comparison and validation with manual segmentation," *Magn. Reson. Med.* **69**(3), 877–883 (2013).
10. M. Zamyadi et al., "Mouse embryonic phenotyping by morphometric analysis of MR images," *Physiol. Genomics* **42**(2), 89–95 (2010).
11. N. Kovacevic et al., "A three-dimensional MRI atlas of the mouse brain with estimates of the average and variability," *Cereb. Cortex* **15**(5), 639–645 (2005).
12. X. J. Chen et al., "Neuroanatomical differences between mouse strains as shown by high-resolution 3D MRI," *NeuroImage* **29**(1), 99–105 (2006).
13. B. Landman and S. Warfield Eds., *MICCAI Workshop on Multi-Atlas Labeling* CreateSpace, Nice, France (2012).
14. A. A. Ali et al., "Automated segmentation of neuroanatomical structures in multispectral MR microscopy of the mouse brain," *NeuroImage* **27**(2), 425–435 (2005).
15. G. Wu, F. Qi, and D. Shen, "Learning-based deformable registration of MR brain images," *IEEE Trans. Med. Imaging* **25**(9), 1145–1157 (2006).
16. J. C. Lau et al., "Longitudinal neuroanatomical changes determined by deformation-based morphometry in a mouse model of Alzheimers disease," *NeuroImage* **42**(1), 19–27 (2008).
17. J. P. Lerch et al., "Automated deformation analysis in the YAC128 Huntington disease mouse model," *NeuroImage* **39**(1), 32–39 (2008).
18. M. H. Bae et al., "Automated segmentation of mouse brain images using extended MRF," *NeuroImage* **46**(3), 717–725 (2009).

19. J. Lee et al., "Evaluation of atlas based mouse brain segmentation," *Proc. SPIE* **7259**, 725943 (2009).
20. J. Nie and D. Shen, "Automated segmentation of mouse brain images using multi-atlas multi-ROI deformation and label fusion," *Neuroinform* **11**(1), 35–45 (2013).
21. A. E. Dorr et al., "High resolution three-dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice," *NeuroImage* **42**(1), 60–69 (2008).
22. G. A. Johnson et al., "Waxholm space: an image-based reference for coordinating mouse brain research," *NeuroImage* **53**(2), 365–372 (2010).
23. Y. Ma et al., "A three-dimensional digital atlas database of the adult C57BL/6J mouse brain by magnetic resonance microscopy," *Neuroscience* **135**(4), 1203–1215 (2005).
24. M. D. Wong et al., "A novel 3D mouse embryo atlas based on micro-CT," *Development* **139**(17), 3248–3256 (2012).
25. D. Adams et al., "Blossomsbury report on mouse embryo phenotyping: recommendations from the IMPC workshop on embryonic lethal screening," *Dis. Model Mech.* **6**(3), 571–579 (2013).
26. T. Zouagui et al., "3D MRI heart segmentation of mouse embryos," *Comput. Biol. Med.* **40**(1), 64–74 (2010).
27. J. I. E. Hoffman and S. Kaplan, "The incidence of congenital heart disease," *J. Am. Coll. Cardiol.* **39**(12), 1890–1900 (2002).
28. S. Roy et al., "Phenotype detection in morphological mutant mice using deformation features," *Lec. Notes Comput. Sci.* **8151**, 437–444 (2013).
29. J. O. Cleary et al., "Cardiac phenotyping in ex vivo murine embryos using μ MRI," *NMR Biomed.* **22**(8), 857–866 (2009).
30. K. Degenhardt et al., "Rapid 3D phenotyping of cardiovascular development in mouse embryos by micro-CT with iodine staining," *Circ. Cardiovasc. Imaging* **3**(3), 314–322 (2010).
31. J. E. Schneider et al., "Rapid identification and 3D reconstruction of complex cardiac malformations in transgenic mouse embryos using fast gradient echo sequence magnetic resonance imaging," *J. Mol. Cell. Cardiol.* **35**(2), 217–222 (2003).
32. W. J. Weninger et al., "Measurements of the diameters of the great arteries and semi-lunar valves of chick and mouse embryos," *J. Microsc.* **234**(2), 173–190 (2009).
33. X. Liang et al., "Towards high-throughput mouse embryonic phenotyping: a novel approach to classifying ventricular septal defects," *Proc. SPIE* **9413**, 94131V (2015).
34. Image Registration Toolkit (IRTK), <https://github.com/BioMedIA/IRTK>
35. D. Mattes et al., "Nonrigid multimodality image registration," *Proc. SPIE* **4322**, 1609–1619 (2001).
36. M. Staring, S. Klein, and J. P. W. Pluim, "A rigidity penalty term for nonrigid registration," *Med. Phys.* **34**(11), 4098–4108 (2007).
37. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vis.* **1**(4), 321–331 (1988).
38. S. C. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(9), 884–900 (1996).
39. P. A. Yushkevich et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *NeuroImage* **31**(3), 1116–1128 (2006).
40. ITK-Snap Medical Image Segmentation Tool, <http://sourceforge.net/projects/itk-snap>
41. M. Tamura et al., "Overdosage of Hand2 causes limb and heart defects in the human chromosomal disorder partial trisomy distal 4q," *Hum. Mol. Genet.* **22**(12), 2471–2481 (2013).
42. Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imaging* **20**(1), 45–57 (2001).
43. S. Joshi et al., "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage* **23**, S151–S160 (2004).

Zhongliu Xie is a senior PhD student at Imperial College London. He received his BSc and MSc degrees in computer science, respectively, from the University of Nottingham in 2010 and Imperial College London in 2011. His current research interests include biomedical image analysis, pattern recognition, computer-aided diagnosis, and optimization algorithms.

Xi Liang is a postdoctoral researcher at IBM Research Australia. She received her BSc degree in computer science and PhD in engineering from the University of Melbourne in 2008 and 2013, respectively. Her current research interests include computer vision, pattern recognition, and medical imaging. She is a member of SPIE.

Liucheng Guo is a senior PhD student at Imperial College London. He received his BSc and MSc degrees in computer science, respectively, from Beijing Institute of Technology in 2009 and Peking University in 2012. His current research interests include optimization algorithms and FPGA-based acceleration.

Asanobu Kitamoto is an associate professor at the National Institute of Informatics, Japan. He received his PhD in electronic engineering from the University of Tokyo in 1997 and joined the National Institute of Informatics. His research focuses on the application of image processing algorithms and image database systems to real-world problems, including bioimage informatics.

Masaru Tamura is a research and development scientist at RIKEN BioResource Center, Japan. He received his PhD in biology from the Graduate University for Advanced Studies (SOKENDAI) in 1998. His research focuses on organogenesis and pattern formation during mouse embryonic development.

Toshihiko Shiroishi is a professor of mammalian genetics and the vice director at the National Institute of Genetics, Japan. He received his PhD in biology from Tohoku University in 1981 and joined the institute in 1984. His research interest is in the genetic control of pattern formation during mouse early embryogenesis. In particular, he focuses on the anteroposterior axis formation in limb development using many mutant mice.

Duncan Gillies is a professor of biomedical data analysis at Imperial College London. He received his BSc degree in engineering science from the University of Cambridge in 1971, his MSc degree in computing from Birkbeck, University of London in 1975, and his PhD in artificial intelligence from Queen Mary University of London in 1977. To date, he has published over 130 papers, been granted five patents, and awarded 24 research grants.