# Phenotype Detection in Morphological Mutant Mice Using Deformation Features

Sharmili Roy[1], Xi Liang[2], Asanobu Kitamoto[3], Masaru Tamura[4], Toshihiko Shiroishi[4], and Michael S. Brown[1]

[1] School of Computing, National University of Singapore
sharmili@comp.nus.edu.sg
[2] National ICT Australia (NICTA), Australia
[3] National Institute of Informatics, Japan
[4] National Institute of Genetics, Japan

**Abstract.** Large-scale global efforts are underway to knockout each of the approximately 25,000 mouse genes and interpret their roles in shaping the mammalian embryo. Given the tremendous amount of data generated by imaging mutated prenatal mice, high-throughput image analysis systems are inevitable to characterize mammalian development and diseases. Current state-of-the-art computational systems offer only differential volumetric analysis of pre-defined anatomical structures between various gene-knockout mice strains. For subtle anatomical phenotypes, embryo phenotyping still relies on the laborious histological techniques that are clearly unsuitable in such big data environment. This paper presents a system that automatically detects known phenotypes and assists in discovering novel phenotypes in $\mu$CT images of mutant mice. Deformation features obtained from non-linear registration of mutant embryo to a normal consensus average image are extracted and analyzed to compute phenotypic and candidate phenotypic areas. The presented system is evaluated using C57BL/10 embryo images. All cases of ventricular septum defect and polydactyly, well-known to be present in this strain, are successfully detected. The system predicts potential phenotypic areas in the liver that are under active histological evaluation for possible phenotype of this mouse line.

## 1 Introduction

Completion of the human genome project brought comprehension of location and sequence of each human gene. Owing to the 99% genetic homology between mouse and human, mouse has been chosen as the principal study model to annotate genetic sequence with its functional information [1], [2]. Gene targeting technology is being actively employed by many international organizations to generate mutant mouse lines by knocking out each of the approximately 25,000 mouse genes (i.e., systematically removing each gene one by one and growing the mouse). High-throughput phenotypic assessment systems are necessary to systematically analyze and interpret the genetic information generated by

K. Mori et al. (Eds.): MICCAI 2013, Part III, LNCS 8151, pp. 437–444, 2013.

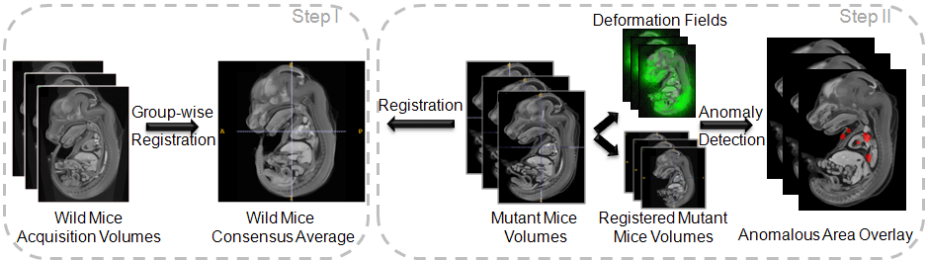



**Fig. 1.** Defect detection consists of two steps. A mean of the normal mouse group is computed in the first step. In the second step, mutant group is registered to the normal mean and the resulting deformations are analyzed to detect defects.

these large-scale mutagenesis programs. A significant proportion of the generated strains are embryonic lethal resulting in the shift towards prenatal phenotyping.

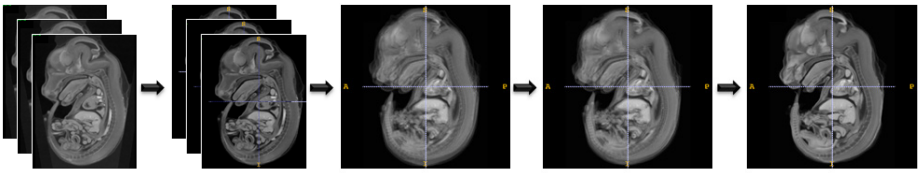
The research community focusses on semi-automatic analysis of anatomical volumetric variation in various mouse strains using representative average images [3], [4], [5]. Although evaluation based on average images may be beneficial for an initial examination, phentoypes that are randomized in position and texture such as the intestines and developing trabeculae of the heart [6] or subtle structural organ failures without large volume changes cannot be characterized using this technique. Another body of work focusses on detection of defects via model-based segmentation [7] or by better data visualization using tissue staining [8], [9]. Segmentation techniques fail if the defect characteristics are unknown or if the anatomy is hard to label such as bone joints. Enhanced visualization, while useful, still requires long expert hours to interpret the data.

In this paper, we present a generalized defect detection framework that automatically computes candidate phenotypic areas without using atlas, segmentation or any defect specific features. Instead, our approach uses deformation fields that are widely used to study anatomical variations [3], [4], [10], [11]. We extract various features from deformation fields obtained by registering mutant mice to a normal mean and combine them to detect coarse, subtle as well as randomized defects (Fig. 1). Statistical characteristics of deformation fields have been previously studied to detect gross defects in mice brain using multi-modality images [11]. Our approach, however, targets a single imaging modality and successfully handles both subtle as well as significantly differing anatomy.

## 2 Methods

### 2.1 Sample Preparation and Imaging Protocol

This study is performed using C57BL/10 mutant mice generated at the National Institute of Genetics, Japan. Post organogenesis, growth and development in the embryo starts at ∼14.0 days post-coitum (dpc). Image registration cannot be applied at stages earlier than this due to unformed or absent organs. Further, at



(a) (b) (c) (d) (e)

**Fig. 2.** This figure illustrates the steps in the computation of normal mean image. (a) Acquisition volume (b) extracted normalized embryo images (c)-(e) consensus average images at rigid, affine and B-Spline registration stages respectively.

a relatively mature stage such as 15.5 dpc accurate registration of abdomen is difficult to achieve due to variation in intestinal position and crowding within the abdominal cavity [4]. Therefore, embryo samples at 14.5 dpc were collected. A total of 14 embryos were used out of which 3 were normal and 11 had chromosomal aberrations. 11 mutant embryos consisted of 3 homozyogotes generated by inbreeding C57BL/10 mice and 8 heterozygotes obtained by cross breeding C57BL/10 and normal littermates. The samples were washed in phosphate buffered saline and fixed in 4% paraformaldehyde until imaging. Before scanning, embryos were soaked in 1 : 3 mixture of lugol solution and double distilled water. Scan was carried out in Scanxmate-E090S 3D $\mu$CT system (Comscantecno, Japan) with the embryos fitted in 1.5 milliliters eppendorf tube fixed using wet paper. Keeping the X-ray source at 60kVp and 130mA, each specimen was rotated 360° in steps of 0.36° generating 1000 projections of $640 \times 480$ pixels. The 3D $\mu$CT data was reconstructed at an isotropic resolution of $9.5 \times 9.5 \times 9.5$ $\mu m^3$.

### 2.2 Normal Mouse Consensus Average Image

Embryo pixels are extracted from the acquisition volumes using Gaussian mixture modeling, thresholding and mathematical morphology (Fig. 2(a) and (b)). Pixel intensity ranges are normalized and a standard group-wise registration routine consisting of rigid, affine and B-Spline registration stages is initiated. Rigid registration corrects differing orientations of individual embryos by choosing a reference and spatially aligning the rest to it. Averaging the rigid registered images results in a blurry reference (Fig. 2(c)) that is not biased towards the geometry of the initial reference because this step does not affect the geometry of the subjects [5]. Embryos are then registered to the blurry reference via affine transformation and the reference is updated (Fig. 2(d)). As a final step B-Spline based non-linear registration is applied to locally align the affine registered embryos to the reference. The non-linear registration is formulated with a similarity energy function comprising of mutual information [12] and a rigidity penalty [13]. 10 iterations of this registration are applied in a multi-resolution fashion where the control point spacing gradually reduces to 8 voxels.

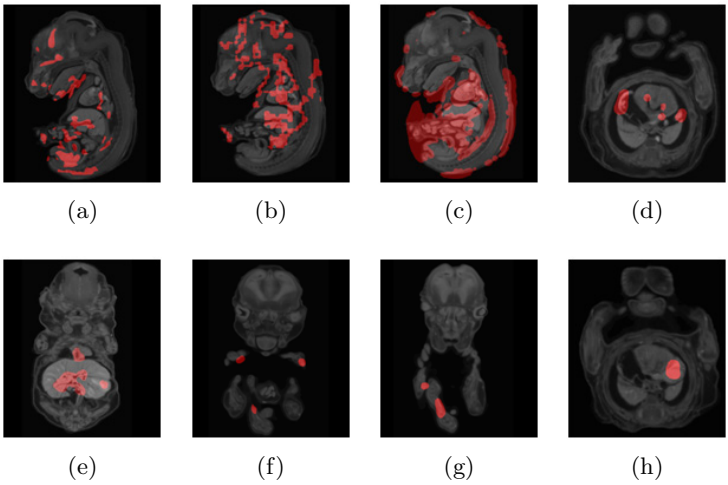



**Fig. 3.** (a),(b),(c) Jacobian, stress and intensity variance masks overlaid on mutant image respectively (d),(e) defective areas identified by $(I_{IV} \cap I_J)$ (f),(g) areas generated by $(I_{IV} \cap I_S)$ (h) defective areas captured by $(I_J \cap I_S)$

The reference image is updated after each iteration leading to the final consensus average (Fig. 2(e)). Elastix toolbox is used to implement this registration scheme [14].

### 2.3 Deformation Features and Masks for Defect Detection

To detect defects in mutant mice, they are registered to the normal average image using the same three-stage registration pipeline as above except that in each stage the reference is always kept fixed to the normal average. Registration of each mouse results in the corresponding deformation field. These deformation fields are used to compute 3D Jacobian maps using determinant of local Jacobian matrix at each voxel [10]. Jacobian determinant greater than one represents voxel expansion and less than one represents voxel compression. Jacobian of deformation is a popular tool to study inter-group structural differences [10], [11]. We apply Jacobian determinant in phenotyping by computing a Jacobian mask, $I_J$, one for each mutant mouse, that selects voxels at which Jacobian determinant is $\delta$ units away from one (Fig. 3(a)). $\delta$ is kept 0.5 for experiments. We realized, however, that Jacobian determinant fails for defects where volume changes are minimal. Further, we find that Jacobian introduces numerous false positives by highlighting areas that are found to be non-defective by the phenotyping experts, thus resulting in low precision. Performance of $I_J$ in detecting known defects namely, Ventricular Septum Defect (VSD) and polydactyly is summarized in Table 1. Detection specificity for both the defects is very low with $I_J$.

To capture subtle defects with low volumetric changes, we compute another deformation feature that we call deformation stress. Deformation stress $(D_s)$ is

**Table 1.** This table compares VSD and polydactyly detection accuracy (in %) of various features. VSD is assumed detected if the ventricular area is highlighted.

| | $I_J$ | $I_S$ | $(I_J \cap I_S)$ $F_1$ | $(I_{IV} \cap I_J)$ $F_2$ | $(I_{IV} \cap I_S)$ $F_3$ | $(F_2 \cup F_3)$ $F_4$ | $(F_1 \cup F_4)$ $F_5$ |
|---|---|---|---|---|---|---|---|
| VSD Sensitivity | 88.8 | 100.0 | 88.8 | 77.7 | 77.7 | 77.7 | 100.0 |
| VSD Specificity | 50.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| Polydactyly Sensitivity | 76.9 | 84.6 | 61.5 | 46.1 | 76.9 | 76.9 | 92.3 |
| Polydactyly Specificity | 48.4 | 80.6 | 87.1 | 90.3 | 87.1 | 87.1 | 87.1 |

computed by dividing the volume into small blocks and measuring the entropy of deformation direction inside each block.

$$D_S(v) = -\sum_{u \in B(v)} p(\theta(u)) \log(p(\theta(u)); \tag{1}$$

$B(v)$ in Equation 1 represents the block in which voxel $v$ lies and $\theta(u)$ is the displacement direction at voxel $u$. For experiments the volume was divided into cuboids of size 8 voxels. Using $D_S$ we compute a mask, $I_S$, that chooses voxel blocks that have high entropy of deformation direction and hence are undergoing incoherent deformation. For experiments $I_S$ selected the top 50% blocks that exhibited high deformation entropy. Fig. 3(b) shows an example of $I_S$ and Table 1 enlists its performance in detecting known phenotypes. Some false positives are introduced due to inclusion of sources and sinks in the deformation field.

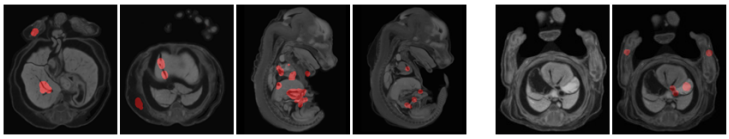
Since $I_J$ and $I_S$ individually fail to detect all defects and both introduce false positives, a simple combination of the two does not give satisfactory results (Table 1). In practice multiple mice from a mouse line are imaged before phenomic analysis is performed. We introduce this group information in defect detection by calculating voxel-wise intensity variance ($V_{IV}$) across the group of mutant mice that are registered to the normal mean.

$$V_{IV}(v) = \frac{1}{N_M - 1} \sum_{i=1}^{N_M} (M_i(v) - N_{Avg}(v))^2. \tag{2}$$

$N_M$, $M_i$ and $N_{Avg}$ in Equation 2 are mutant mouse population size, $i^{th}$ registered mutant mouse image and the normal consensus average respectively. From $V_{IV}$ we compute a mask $I_{IV}$ by selecting the top 50% voxels that have high intensity variance and hence low registration accuracy(Fig. 3(c)).

We find that many false positives introduced by $I_J$ and $I_S$ are pruned when these features are combined with $I_{IV}$. Table 1 lists the accuracies when the detection criterion is $(I_{IV} \cap I_J)$, $(I_{IV} \cap I_S)$ or both combined. From experiments on C57BL/10 mutant mice we find that $(I_{IV} \cap I_J)$ mainly captures VSD and unusually wide liver lobe junctions. Figs. 3(d) and 3(e) show some detection results obtained by this factor. $(I_{IV} \cap I_S)$ captures polydactyly and unnatural position and deformations of tail and limbs (Figs. 3(f) and 3(g)).

It is possible to have inaccurately registered morphological structures that are not selected by the intensity variance mask. Uniform body cavities (dark



(a) (b)

**Fig. 4.** (a) Defect detection results in the liver lobe junctions, heart and intestine of C57BL/10 mice (b) the left and right images depict a healthy heart and the misjudged defect respectively

regions) or muscular organs (like liver lobes and heart atria) are some examples where we find that intensity variance is low due to spatially uniform intensity values even though there is a abnormality. ($I_J \cap I_S$) addresses regions where both Jacobian and stress are high irrespective of the intensity variance. By adding this term in the detection rule we are able to detect spatially uniform defective regions. Some secondary phenotypes like enlarged heart atrium due to high blood pressure induced by VSD are captured by this term as shown in Fig. 3(h).

Combining the three terms we propose the defect detection rule as

$$I_{Defect} = (I_{IV} \cap I_J) \cup (I_{IV} \cap I_S) \cup (I_J \cap I_S). \tag{3}$$

Table 1 enlists the performance of this detection rule. Number of detected regions can be readily increased or decreased by relaxing or tightening the thresholds while generating the Jacobian, deformation stress and intensity variance masks. Simple morphological operations like dilation and erosion are applied as noise reduction measures to clean up the detection results.

## 3 Results

Complete phenomic analysis of a mouse strain is a very tedious and slow process. C57BL/10 strain is still under investigation and hence full phenotypic characteristic consisting of all phenotypic defects is yet unknown. Therefore, even though we can evaluate the detection rule (Equation 3) in terms of precision, a formal evaluation of recall is not possible. VSD and polydactyly are two established genetic defects in C57BL/10 mice. We compute sensitivity and specificity of the detection algorithm with respect to these two defects.

When evaluated over the mutant database of 3 homozygote and 8 heterozygote embryos, the algorithm detected all cases of VSD without generating any false positives. Out of the 13 cases of polydactyly 12 were successfully detected and 1 was missed. The missed case belonged to the only mouse in the database that had its umbilical chord removed resulting in registration errors at the nearby areas. 4 false positive polydactyl cases were reported in situations where due to high proximity toes of both the feet seemed fused in the 3D renderings.

To evaluate the rest of the detected areas, a user study was conducted with phenotyping experts having long experience in mouse imaging and phenotyping.

The expert comments were very encouraging and they noted that all the regions detected by the algorithm had biological significance. Some of the areas were due to genetic defects, some due to non-genetic defects and some due to organogenesis or procedural interventions. After careful histological examination, it was found that majority (57%) of the total detected regions belonged to genetic defects. Apart from known phenotypes, the algorithm detected areas in the liver lobe junctions that are candidate areas for potential phenotype of this mouse strain and are under active phenotypic evaluation.

14% output regions were detected due to malformed body cavities. Though these regions do not represent defects due to genetic makeup, they still signify biological malformations. Another 8% regions were noted to be due to genesis and extensive developmental remodeling of gonads at this gestational stage. The rest of the output was attributed to blood clots, randomized umbilical chord regions and pancreatic genesis.

When the detection rule was applied to wild-type mice, some areas were reported. These areas represent blood clots, umbilical chord, malformed body cavities and pancreatic and gonadic organogenesis. One false positive was generated for heart septal defect in a case where low spatial intensity variance makes the judgement hard even for an unexperienced human eye (Fig. 4(b)). With further image processing it is possible to improve the detection accuracy by neglecting the high intensity blood clots and masking out umbilical chord regions.

## 4 Discussion and Conclusion

We have presented a generic deformation based defect detection framework for 3D $\mu$CT images of mutant mice. Our system has the potential to greatly enhance phenotyping throughput by automatically detecting all known phenotypes. Unlike other algorithms designed to detect specific known defects, our system also highlights candidate novel defects that may not be readily recognized by human experts due to absence of significant visual features. Owing to voxel-by-voxel analysis, defects are localized to sub-structures and those affecting multiple structures are visualized collectively. Though our evaluation database is small, the results clearly establish the potential of the proposed system in patterning defects. Our framework can be easily adapted to examine other 3D scan images amenable to registration. We acknowledge that the registration method may effect the detection results, however, the registration scheme used in the paper is widely employed in mice phenotyping [3], [4], [5]. Deformation field resulting from only the non-linear registration step is used for defect detection. The detection performance is found to be fairly robust to parameter variation.

Since the proposed framework is independent of the defect features, classification of defects into those that are genetically induced and those that are not is out of scope for the current system. Currently we provide frequency of occurrence as an indicator of whether or not a defect is genetic. As an example, since VSD and polydactyly are detected in all homozygote embryos, the probability of these defects being genetic is reported to be 100%. Similar probabilities

are assigned to all detected regions. In future we plan to use advanced image processing and statistical techniques to device classifiers that can differentiate between genetic and non-genetic defects.

## References


1. Collins, F.S., Rossant, J., Wurst, W.: A mouse for all reasons. Cell 128, 9–13 (2007)
2. Mouse Genome Sequencing Consortium: Initial Sequencing and comparative analysis of the mouse genome. Nature 420, 520–562 (2002)
3. Zamyadi, M., Baghdadi, L., Lerch, J.P., Bhattacharya, S., Schneider, J.E., Henkelman, R.M.: Mouse embryonic phenotyping by morphometric analysis of MR images. Physiol. Genomics 42A, 89–95 (2010)
4. Wong, M.D., Dorr, A.E., Walls, J.R., Lerch, J.P., Henkelman, R.M.: A novel 3D mouse embryo atlas based on micro-CT. Development 139(17), 3248–3256 (2012)
5. Cleary, J.O., Modat, M., Norris, F.C., Price, A.N., Jayakody, S.A., Martinez-Barbera, J.P., Greene, N.D.E., Hawkes, D.J., Ordidge, R.J., Scambler, P.J., Ourselin, S., Lythgoe, M.F.: Magnetic resonance virtual histology for embryos: 3D atlases for automated high-throughput phenotyping. Neuroimage 54(2), 769–778 (2011)
6. Nieman, B.J., Wong, M.D., Henkelman, R.M.: Genes into geometry: imaging for mouse development in 3D. Curr. Opin. Genet. Dev. 21(5), 638–646 (2011)
7. Norris, F.C., Modat, M., Cleary, J.O., Price, A.N., McCue, K., Scambler, P.J., Ourselin, S., Lythgoe, M.F.: 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)
8. Degenhardt, K., Wright, A.C., Horng, D., Padmanabhan, A., Epstein, J.A.: Rapid 3D phenotyping of cardiovascular development in mouse embryos by micro-CT with iodine staining. Circ. Cardiovascular Imaging 3(3), 314–322 (2010)
9. Cleary, J.O., Price, A.N., Thomas, D.L., Scambler, P.J., Kyriakopoulou, V., McCue, K., Schneider, J.E., Ordidge, R.J., Lythgoe, M.F.: Cardiac phenotyping in ex vivo murine embryos using $\mu$MRI. NMR Biomed. 22(8), 857–866 (2009)
10. Xie, Z., Yang, D., Stephenson, D., Morton, D., Hicks, C., Brown, T., Bocan, T.: Characterizing the regional structural difference of the brain between tau transgenic (rTg4510) and Wild-Type Mice using MRI. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 308–315. Springer, Heidelberg (2010)
11. Nieman, B.J., Flenniken, A.M., Adamson, S.L., Henkelman, R.M., Sled, J.G.: Anatomical phenotyping in the brain and skull of a mutant mouse by magnetic resonance imaging and computed tomography. Physiol. Genomics 24(2), 154–162 (2006)
12. Mattes, D., Haynor, D.R., Vesselle, H., Lewellyn, T.K., Eubank, W.: Nonrigid multimodality image registration. In: Proc. SPIE, vol. 4322, pp. 1609–1619 (2001)
13. Staring, M., Klein, S., Pluim, J.P.W.: A rigidty penalty term for nonrigid registration. Med. Phys. 34(11), 4098–4108 (2007)
14. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W.: elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging 29(1), 196–205 (2010)