Reproducible Evaluation of Registration Algorithms for Movement Correction in Dynamic Contrast Enhancing Magnetic Resonance Imaging for Breast Cancer Diagnosis

I. A. Illan[1,2]([✉]), J. Ramirez[2], J. M. Gorriz[2], K. Pinker[3,4],
and A. Meyer-Baese[1]

[1] Scientific Computing Department, Florida State University, Tallahassee, USA
[2] Department of Signal Theory, Networking and Communications,
DaSCI (Data Science and Computational Intelligence Research Institute),
Universidad de Granada, Granada, Spain
illan@ugr.es
[3] Department of Radiology, Memorial Sloan-Kettering Cancer Center,
New York, USA
[4] Department of Biomedical Imaging and Image-guided Therapy,
Division of Molecular and Gender Imaging,
Medical University of Vienna/AKH Wien, Vienna, Austria

Abstract. Accurate methods for computer aided diagnosis of breast cancer increase accuracy of detection and provide support to physicians in detecting challenging cases. In dynamic contrast enhancing magnetic resonance imaging (DCE-MRI), motion artifacts can appear as a result of patient displacements. Non-linear deformation algorithms for breast image registration provide with a solution to the correspondence problem in contrast with affine models. In this study we evaluate 3 popular non-linear registration algorithms: MIRTK, Demons, SyN Ants, and compare to the affine baseline. We propose automatic measures for reproducible evaluation on the DCE-MRI breast-diagnosis TCIA-database, based on edge detection and clustering algorithms, and provide a rank of the methods according to these measures.

Keywords: Medical image processing · Reproducibility · DCE-MRI
Registration · Diffeomorphism · Optical flow · Non-affine registration

1 Introduction

Early and accurate detection is a key factor in maximizing probability of survival of breast cancer, the second most common cause of cancer death, after lung cancer [17]. Computer aided diagnosis (CAD) systems support the physicians

in decision making by means of imaging processing techniques. Dynamic contrast enhancing magnetic resonance imaging (DCE-MRI) allows to obtain non-invasive information about tissue dynamics that allows for cancer identification, and CAD systems have been developed to assist clinicians in the analysis. Accuracy of such CADs depend on the preprocessing steps such as registration. Registration is a challenging task due to the highly deformable nature of the breast, and linear methods are unsuitable.

Nowadays, with the impetuous advance of data-sharing and code-sharing in science, reproducibility has become a key factor in new developments and analysis. Medical image processing traditionally has been focused in giving a solution to the registration problem, together with the segmentation problem. In the field of DCE-MRI of the breast, there has been considerable interest in non-linear deformation modeling to solve the correspondence problem due to the deformable nature of the breast. Although successful methods for enhancement subtraction have been developed [11], the non-rigid registration step in DCE-MRI is often performed using state-of-the art approaches, such as diffeomorphic based ones, or partial-differential-equations-based ones [6,7,13,18]. However, there is no clear argument for a preferable choice nor a systematic comparison of the proposed solutions. Led by the neuroimaging community, a large variety of high quality software tools, projects and analysis are publicly available to the medical imaging community that solve the registration problem. Concretely, consistent effort has been dedicated in the last decade to evaluate registration algorithms, such as with the projects the mindboggle (http://www.mindboggle.info) and NIREPS (http://www.nirep.org/links).

Given the importance and challenges in nonlinear deformation algorithms for breast image registration, a systematic analysis with reproducible results can provide an evaluation of the available registration tools and its efficacy in the aforementioned task, setting a baseline for ulterior developments. Existing evaluations of registration methods are often based on landmarks. Landmarks require some expert intervention, and are prone to subjectivity and non-reproducibility. We propose an automatic method for registration evaluation based on an algorithmic-driven extraction of morphological features: contours and volume. For the extraction of contours we propose the use of the Canny edges detection algorithm and k-means algorithm to extract tissue-specific volumes. We evaluate 3 popular non-linear registration algorithms: MIRTK, Demons, SyN Ants, and compare to the affine baseline. We use a publicly available DCE-MRI breast database and rank the tested registration algorithms according to the measures proposed.

2   Methods

2.1   TCIA Database

The data used in this work is the Breast-Diagnosis collection [3] from the cancer imaging archive (TCIA)[4]. The Breast-Diagnosis collection contains cases that are high-risk normals, DCIS, fibroids and lobular carcinomas. Each case has 3

or more distinct MR pulse sequences from a Phillips 1.5 T, such as T2, STIR and BLISS. For the DCE-MRI, the volume of Magnevist (Bayer) gadolinium contrast injected into the brachial vein is based on a rule of thumb which in ml's is 10% of the patient weight. The injection itself is 6 or 7 s, at a rate of 3cc per second. The first dynamic sequence is started 1 min after the injection is started.

2.2   The Registration Problem

One of the main medical image processing problems is the registration problem. The goal of the registration is to find the transformation T that maps each point x of a fixed image F to a point y = T(x) in image moving image M, so that some measure of similarity $E(F, M)$ between F and M is minimized. In the affine case, this transformation can be parametrized by a small set of parameters (12 in 3D). When non-rigid transformations are allowed, a non-parametric characterization of the transformation must be given, usually in terms of a displacement field.

This study will consider well established deformable techniques for deformable registration: Diffeomorfic Demons, MIRTK, SyN [8]:

- **Diffeomorphic Demons** is a non-parametric algorithm based in optical flow theory that generalizes Thirion's Demons algorithm with a diffeomorphic spatial transformation [20]. This method alternates between the computation of warping forces and smoothing. The Demons algorithm may be related to a Taylor expansion of the squared difference between the fixed and moving image, with some regularization in the form of fluid-like equations.
- **MIRTK** [15,16] uses a combined transformation T which consists of a global affine transformation and a local transformation. The local transformation describes any local deformation required to match the anatomies of the subjects using a free-form deformation (FFD) model based on B-splines. The basic idea of FFDs is to deform an object by manipulating an underlying mesh of control points. The resulting deformation controls the shape of the 3-D object and can be written as the 3-D tensor product of the familiar 1-D cubic B-splines. The lattice of control points is defined as a grid with uniform spacing which is placed on the underlying reference image. The optimal transformation is found using a gradient descent minimization of a cost function associated with the global transformation parameters as well as the local transformation parameters. The cost function comprises two competing goals: The first term represents the cost associated with the voxel-based similarity measure, in this case normalised mutual information, while the second term corresponds to a regularization term which constrains the transformation to be smooth.
- **The symmetric normalization (SyN)** methodology uses a symmetric parameterization of the shortest path of diffeomorphisms connecting two anatomical configurations [2]. The SyN formulation uses a bidirectional gradient descent optimization which gives results that are unbiased with respect to the input images. SyN also provides forward and inverse continuum mappings that are consistent within the discrete domain and enables both large

and subtle deformations to be captured. Specific performance characteristics depend upon the range of similarity metrics chosen for the study and the velocity field regularization.

2.3   Automatic Contour and Volume Estimation

For evaluation purposes, two features are extracted from the 3D MRI images: contours and volumes. Contours are extracted by a Canny edge detector algorithm, that provides the external surface as well as interior structures. The Canny filter is a multi-stage edge detector that uses a derivative-of-a-Gaussian in order to compute the intensity of the gradients. We used the simpleITK [10,21] implementation and set the variance of the Gaussian $[\sigma_x, \sigma_y, \sigma_z]$ to $[25, 25, 25]$. The Canny edge detector has been reported to have superior capabilities as other edge detectors as Sobel [14].

In brain imaging, registration evaluation is usually performed by evaluating the overlap between different brain structures, as the thalamus or the hippocampus, defined in some standard atlas. The differences on breast density, shape and randomness of breast lesion locations makes the definition of a breast atlas inviable. Therefore, the volume overlap evaluation of the registration algorithm lacks of a natural volume subdivision of the breast to compare to. Consequently, subvolumes and structures must be either manually defined case by case, a task prone to inaccuracies and subjectivity, or extracted automatically from a clustering algorithm. Here, different breast structures are extracted by the k-means algorithm. Since the number of different tissues is expected to be low, following Thirion et al. [19], the k-means algorithm performs better than other competitive alternatives, as the Ward agglomerative clustering with spatial constraints. We use the k-means implementation from nilearn [1] with standardization and smoothing (FWHM = 10) and 3 clusters.

3   Results

For file preparation, the first step is convert the DICOM files to NIFTI format. All the tested algorithms require NIFTI files as input, since it is the standard format in neuroimaging. This task is performed with the use of dcm2niix [9]. Next, the images are masked to remove the strong signals from internal organs. It is important for a registration algorithm to remove the signals from interior cavities of the chest. In the case of non-affine transformations, the variability of these regions produces the highest metric values on those localized regions, thus dominating in the optimization algorithm. The mask is defined as the middle plane in the axial direction, and all the voxels inside the plane are set to background.

The masked images are registered to the pre-contrast images using the different tested algorithms. Contours and volumes are extracted from each sequence and two metrics are employed to evaluate the overlap between solutions:

– **Dice similarity score (DCS)** Also known as F1-score, the DSC evaluates the overlap between the true labels in the fixed image (F) and the test labels in the moving image (M) by:

$$DSC(F, M) = \frac{2|F \cap M|}{|F| + |M|} \tag{1}$$

– **Jaccard similarity score (JCS)** Evaluates the overlap between the fixed image and the moving image with respect to the whole volume of both.

$$JSC(F, M) = \frac{|F \cap M|}{|F \cup M|} = \frac{|F \cap M|}{|F| + |M| - |F \cap M|} \tag{2}$$

A Welch t-test is performed between the values obtained by each algorithm and the best performance, in order to detect significant differences. Figure 1 shows a boxplot of the evaluation parameters on different algorithms, while Fig. 2 shows the adequacy of the DSC parameter to evaluate the registration performance: in time point 5 a movement is registered and not corrected by the affine transformation, while the deformation transformation minimizes the differences in subsequent times in the sequence (Table 1).
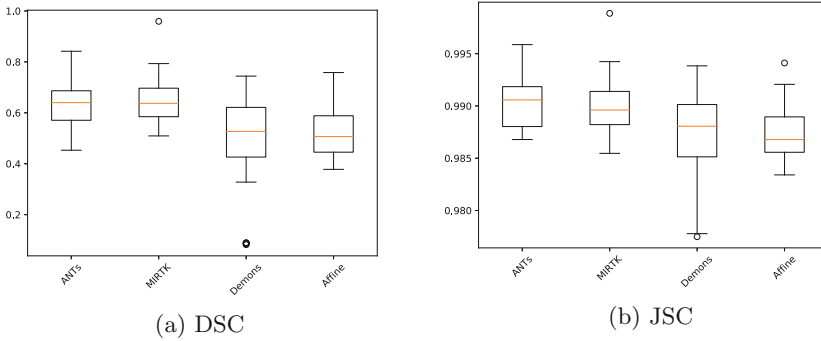

(a) DSC   (b) JSC

Fig. 1. DSC and Jaccard scores for contour overlap evaluation

Table 1. Average performance parameters on validating data for contour and volume

	Contour		Volume
	DSC	JSC	DSC
Affine	0.517 ± 0.007	0.9873 ± 0.0001	0.975 ± 0.001
ANTs	0.631 ± 0.009	0.9904 ± 0.0001	0.979 ± 0.001
MIRTK	0.646 ± 0.007	0.9898 ± 0.0001	0.982 ± 0.006
Diffeomorphic Demons	0.497 ± 0.029	0.987 ± 0.001	0.965 ± 0.001

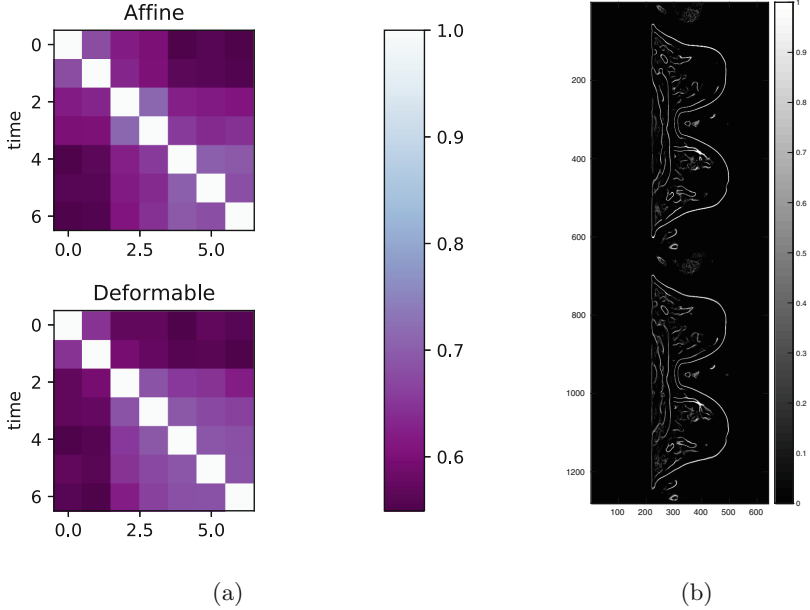(a)   (b)

Fig. 2. (a) DSC comparison of an example affine and deformable registration for 6 timepoints in the DCE-MRI sequence. Pre and post contrast differences can be noted, as well as post-contrast motion correction with the deformable model. (b) Center axial slice of mean contour overlap after deformable registration (up) and affine registration (down). Increase in sharpness in chest-wall lines can be noted in deformable-registration mean contours.

4   Discussion

The results presented here combine two independent measures of 3 registration algorithms. Taking altogether, diffeomorphic Demons appear to be the registration algorithm with less consistent performance. In the contour evaluation, ANTs and MIRTK provide similar performance. Although MIRTK reaches the maximum DSC, ANTs has the maximum JSC. This fact can be explained from the effect on chest-wall contours, where ANTs algorithm produces larger deformations to remove dissimilarities, not related to motion artifacts. When perfoming a Welch t-test, the null hypothesis of having equal values for ANTs and MIRTK can not be rejected (p-value 0.32), while is rejected when comparing with the other methods. Regarding the volumetric measures, it is remarkable that independent measures show a consistent behavior, suggesting that this results may generalize to other databases. These results are in coherence with other performance studies involving some of the studied algorithms [5,8].

For registration evaluation its implicitly assumed in this work that within subject differences due to enhancing tissues do not modify substantially the contour and volume of organic structures. However, one of the limitations of the

present study relies on the automatic algorithms employed to obtain the structure measures, as they are affected by the intensity changes due to enhancement. To alleviate this effect, results within pre-contrast images are studied independently from post-contrast ones. It has been argued that the intensity enhancement of MRI signals has itself an effect on registration algorithms [12]. From the results presented in this work, it can be argued that the effect of intensity enhancement on the registration algorithms is not as dramatic as the effect of interior chest organs, as preliminary results in this study showed. Once the masking preprocessing step is done, the registration algorithms seem to depend weakly on the enhancement, as the difference between contour and volume measures display. A future line of research will be to define an enhancing-independent measure of structure features in breast MRI to quantitatively evaluate the enhancing effect on registration algorithms.

5   Conclusion

We have presented a reproducible analysis on registration performance in breast DCE-MRI for 3 non-rigid deformation algorithms on a TCIA open dataset. Two automatic measures are calculated containing information on contours and volume overlap, and manual intervention is avoided. Results suggest that ANTs and MIRTK should be between the preferable choices for registration, and this result may generalize to other datasets.

References

1. Abraham, A.: Machine learning for neuroimaging with scikit-learn. Front. Neuroinform. **8**, 1–14 (2014)
2. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Image Anal. **12**(1), 26–41 (2008)
3. Bloch, Nicolas, B., Jain, Ashali, Jaffe, C.C.: Data from breast-diagnosis. The Cancer Imaging Archive (2015)
4. Clark, K., et al.: The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging **26**(6), 1045–1057 (2013)
5. Diez, Y., et al.: Comparison of methods for current-to-prior registration of breast DCE-MRI. In: Fujita, H., Hara, T., Muramatsu, C. (eds.) IWDM 2014. LNCS, vol. 8539, pp. 689–695. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07887-8_95
6. Ebrahimi, M., Martel, A.L.: A general PDE-framework for registration of contrast enhanced images. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5761, pp. 811–819. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04268-3_100

7. Kim, M., Wu, G., Shen, D.: Hierarchical alignment of breast DCE-MR images by groupwise registration and robust feature matching. Med. Phys. **39**(1), 353–366 (2012)
8. Klein, A., et al.: Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage **46**(3), 786–802 (2009)
9. Li, X., Morgan, P.S., Ashburner, J., Smith, J., Rorden, C.: The first step for neuroimaging data analysis: DICOM to NIfTI conversion. J. Neurosci. Methods **264**, 47–56 (2016)
10. Lowekamp, B.C., Chen, D.T., Ibanez, L., Blezek, D.: The design of SimpleITK. Front. Neuroinform. **7**, 1–45 (2013)
11. Martel, A.L., Froh, M.S., Brock, K.K., Plewes, D.B., Barber, D.C.: Evaluating an optical-flow-based registration algorithm for contrast-enhanced magnetic resonance imaging of the breast. Phys. Med. Biol. **52**(13), 3803 (2007)
12. Melbourne, A., Atkinson, D., White, M.J., Collins, D., Leach, M., Hawkes, D.: Registration of dynamic contrast-enhanced MRI using a progressive principal component registration (PPCR). Phys. Med. Biol. **52**(17), 5147 (2007)
13. Ong, R.E., Ou, J.J., Miga, M.I.: Non-rigid registration of breast surfaces using the laplace and diffusion equations. BioMed. Eng. OnLine **9**, 8 (2010)
14. Othman, Z., Haron, H., Kadir, M.R.A., Rafiq, M.: Comparison of Canny and Sobel edge detection in MRI images. Comput. Sci. Biomech. Tissue Eng. Group Inf. Syst. 133–136 (2009)
15. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging **18**(8), 712–721 (1999)
16. Schnabel, J.A., et al.: Validation of nonrigid image registration using finite-element methods: application to breast MR images. IEEE Trans. Med. Imaging **22**(2), 238–247 (2003)
17. Siegel, R.L., Miller, K.D.: Cancer statistics. CA Cancer J. Clin. **68**(1), 7–30 (2018)
18. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: a survey. IEEE Trans. Med. Imaging **32**(7), 1153–1190 (2013)
19. Thirion, B., Varoquaux, G., Dohmatob, E., Poline, J.B.: Which fMRI clustering gives good brain parcellations? Front. Neurosci. **8**, 1–13 (2014)
20. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic demons: efficient non-parametric image registration. NeuroImage **45**(1, Suppl. 1), S61–S72 (2009)
21. Yaniv, Z., Lowekamp, B.C., Johnson, H.J., Beare, R.: SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research. J. Dig. Imaging **31**(3), 290–303 (2018)