INVNET: A DEEP LEARNING APPROACH TO INVERT COMPLEX DEFORMATION FIELDS

Marek Wodzinski[1] *and Henning Müller*[2]

[1] AGH University of Science and Technology
Department of Measurement and Electronics, Krakow, Poland
wodzinski@agh.edu.pl
[2] University of Applied Sciences Western Switzerland
Information Systems Institute, Sierre, Switzerland
henning.mueller@hevs.ch

ABSTRACT

Inverting a deformation field is a crucial part for numerous image registration methods and has an important impact on the final registration results. There are methods that work well for small and relatively simple deformations. However, a problem arises when the deformation field consists of complex and large deformations, potentially including folding. For such cases, the state-of-the-art methods fail and the inversion results are unpredictable. In this article, we propose a deep network using the encoder-decoder architecture to improve the inverse calculation. The network is trained using deformations randomly generated using various transformation models and their compositions, with a symmetric inverse consistency error as the cost function. The results are validated using synthetic deformations resembling real ones, as well as deformation fields calculated during registration of real histology data. We show that the proposed method provides an approximate inverse with a lower error than the current state-of-the-art methods.

***Index Terms*—** Image registration, Deformation field, Deep learning, Missing data

1. INTRODUCTION

The problem of finding an inverse deformation field such that $u \circ u^{-1} = \mathrm{Id}$ is a crucial problem for numerous image registration algorithms, especially considering methods focused on preserving the registration symmetry. For these problems, registering source to target should calculate a deformation field that is an exact inverse of the deformation field calculated during target to source registration. The quality of these algorithms strongly depends on the symmetry preservation during the registration process, e.g. for the registration of the spinal cord [1]. On the other hand, the problem is also important for algorithms where the real deformation field is inherently non-diffeomorphic and thus non-invertible [2]. Even for such cases, it can be useful to find an approximate inverse mapping with the best possible inverse consistency, e.g. to calculate dose margins during radiotherapy planning. Current state-of-the-art methods do not handle the regions where a Jacobian determinant is not positive well. The behavior is usually not well-defined and the resulting inverse consistency error is high.

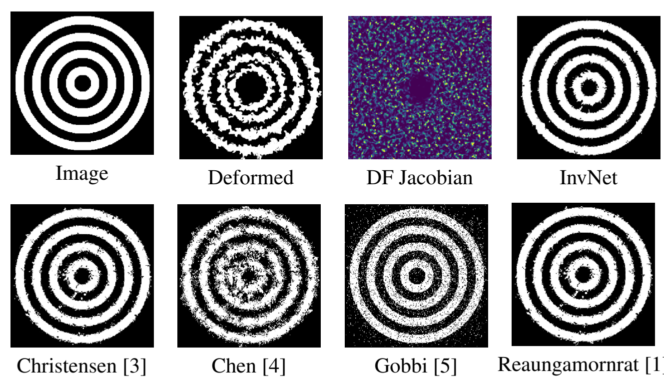

Fig. 1. An example visualization of an image warped with the generated, synthetic deformation field and the same image warped with $u \circ v$, where v is the output of the given inversion method.

Related work: The majority of the state-of-the-art methods for inverting deformation fields were proposed together with new image registration algorithms. One of the most mature methods was proposed in [3] together with a method for consistent image registration. As will be shown in the results section, the method is currently the most stable and it provides the most accurate results, even for non-invertible deformation fields. In [5], the authors proposed another approach to find the inverse of general nonlinear transformations based on Newton's method. The implementation of the method was introduced as a default inversion algorithm in the VTK (Visualization ToolKit) library. The algorithm iteratively updates the result using a smart initial guess and an explicit Jacobian calculation. A dedicated, stable method rooted in fixed-point theory, with a proven convergence when the Lipschitz condition is met, was proposed in [4]. The authors showed promising results using both synthetic and real deformation fields and provided a good theoretical explanation. The most recent algorithm was proposed together with a symmetric, stable and accurate multimodal registration algorithm [1], as an algorithm to find an inversion of a given diffeomorphism. The algorithm works well for simple and invertible deformations. However, the error for more complex cases, including folding, remains hard to estimate.

There are several articles discussing similar topics including invertible neural networks [6], deep regularizers together with a solid theoretical formulation [7], general inverse ill-posed problems in imaging, with the image reconstruction as an example [8], or the use of symmetric properties to directly improve the registration [9, 10, 11].

Contribution: In this work, we propose a dedicated deep network based on the U-Net encoder-decoder architecture to improve the calculation of the inverse mapping of the deformation field. The most important novelty of our approach is that the network is trained using the symmetric inverse consistency as a cost function. As a result, the proposed method attempts to provide the lowest inverse consistency error, even for non-invertible deformation fields with no correct one-to-one inverse mapping. Such a method can be useful for image registration with partial or missing data. It is noteworthy that in a real-life scenario, the majority of the deformations are non-invertible because the problem of missing data is ubiquitous (e.g. due to weight loss, tumor resection, surgical instrument insertion, a non-overlapping region of interest and many other reasons). Since a deformation field itself cannot fully remove an object or add new non-existing data, the exact inversions for such cases do not exist and the best possible approximation is useful.

2. METHODS

The purpose of the proposed method is a regularization of an initially inverted deformation field in a way that minimizes the inverse consistency error. It is important to emphasize that the method does not attempt to calculate a direct inverse mapping. Even though we experimentally verified that this is possible, the network would require many more parameters and the input patch had to be considerably smaller, resulting in a much longer computation time. Since the problem is solved for deformation fields with a positive Jacobian, we decided that it is a better idea to propose a network that just improves the approximate solution in a way that minimizes the symmetric inverse consistency.

The overall algorithm and network structure are presented in Figure 2. The solution consists of two parts, a deterministic, direct inversion method based on a slightly modified algorithm described in [3], and a deep network that takes as an input the result of the approximate inversion. We verified

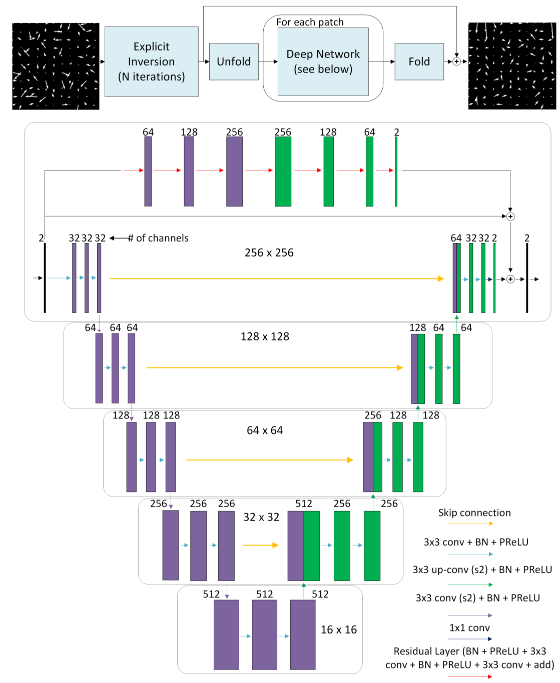

Fig. 2. Visualization of the proposed method and network architecture. For residual layers the interconnections are omitted for presentation clarity.

that running just a few (in our case 5) iterations of the direct inversion provides better final results. The network itself also consists of two separate parts. The first one is based on the slightly modified U-Net architecture dedicated to reconstruction [8] and the second one is a residual encoder-decoder being able to learn the identity mappings [12]. The motivation behind the U-Net branch is rather obvious, a general and strong feature encoding/decoding. However, the motivation behind the residual encoder-decoder deserves a more detailed explanation. The discussed problem is an example of a regression task, similar to image reconstruction or image registration. Therefore, one of the big challenges was to ensure that the network is capable of preserving fine details. Instead of just attempting to find the inverse, we were constantly ensuring that the network is able to learn the identity transformation (just the network, not the whole structure presented in Figure 2). In addition to the residual branch we replaced all pooling operations with strided convolutions. The network operates at the patch-level. The images are unfolded into overlapping patches, inverted and then composed again using the patch centers with a stride equal to 25% of the patch size.

The cost function was defined as follows:

$$J(u, v) = \frac{||u \circ v - \mathrm{Id}|| + ||v \circ u - \mathrm{Id}||}{2}, \qquad (1)$$

where u is the input deformation field, v is the calculated approximate inverse mapping, Id is the identity transformation and ∘ denotes the composition of the deformation field. We decided to use the symmetric formulation instead of a single composition because we verified that such an approach directly regularizes the network and minimizes the influence of boundary conditions. As a result, the network converged to a more accurate solution.

The network was trained using synthetic deformation fields. The deformations were generated randomly using random Gaussian fields, a B-Spline transformation model, affine transformations, local contracting or expanding transformations and various other approaches. The parameters of all transformations were generated randomly each time and the transformations were composed in random order with a given probability. This resulted in a practically infinite data set size, which made it impossible to observe any overfitting. A similar approach can be used to pre-train deep networks dedicated to image registration. Both the validation and the test sets were generated using substantially different parameters and distributions to validate that the method has the generalization ability.

3. RESULTS

The proposed method was validated using deformations that can be divided into three subdomains. The first one consists of deformations with very large magnitude where the Jacobian determinant is very close to zero but is rarely negative. Such deformations can often be seen in practice for image registration algorithms where the initial global alignment cannot be explicitly separated from the subsequent nonrigid registration. The latter includes deformations smaller in magnitude but resembling tumor resection, the insertion of a surgical instrument or other sources of missing or redundant data. These deformations are inherently non-invertible and the Jacobian determinant happens to be negative. Finally, we calculated deformation fields using state-of-the-art affine/nonrigid registration methods on an open, medical dataset [13, 14]. An exemplary visualization of the applied deformations on a synthetic image is shown in Figure 1. The mean inverse consistency errors for the proposed method and state-of-the-art methods are shown in Figure 3, Figure 4, and Figure 5. All methods used for comparison were run until convergence. The proposed method noticeably reduced the mean inverse consistency error, especially for the non-invertible deformation fields. It is interesting that not only negative Jacobians are challenging but also non-smooth regions with a very large, positive determinant value.

4. DISCUSSION AND CONCLUSION

The results show that the proposed method calculates the approximate inverse with lower mean inverse consistency error compared to the currently applied state-of-the-art methods. However, there is still much to improve. First, the forward

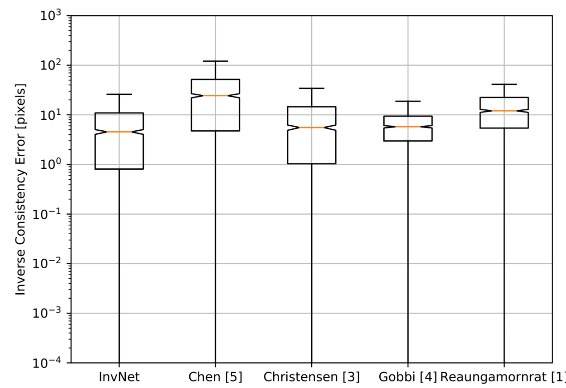

Fig. 3. Inverse consistency error for large and complex synthetic deformations with Jacobian being rarely negative (but with regions with very large or small values).

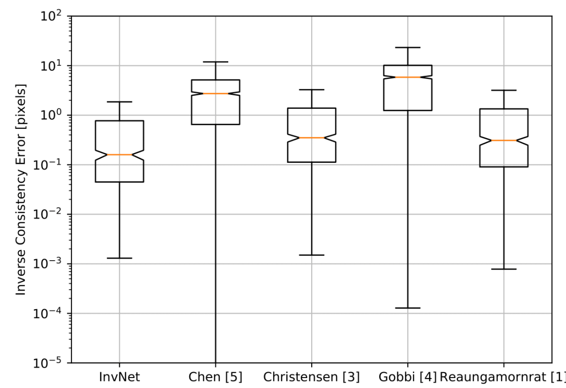

Fig. 4. Inverse consistency error for relatively small synthetic deformations but with zero or negative Jacobian resembling missing or partial data problems.

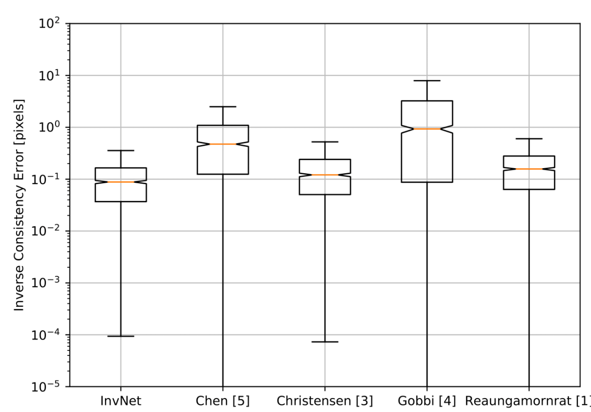

Fig. 5. Inverse consistency error for deformations calculated by state-of-the-art image registration algorithms using an open medical dataset.

pass takes more time than dozens of iterations of the remaining methods (all the methods are implemented on a GPU using PyTorch). The forward pass time can be reduced by replacing the residual branch with a faster alternative, possibly including smarter, direct connections in the U-Net part of the network. Second, an initial guess about the regions to regularize can be beneficial. The Jacobian determinant can be calculated before the forward pass to choose only the image patches where it is not positive. Last, we think that it is possible to propose a better cost function than the symmetric inverse consistency error. Even though it can be seen as a natural choice, the convergence is much slower compared to e.g. fully supervised methods with a known ground-truth.

Noteworthy, the method is not dedicated to simple diffeomorphisms. For such cases the problem is well-defined and the state-of-the-art methods can easily find the exact inverse (the inverse consistency error at the level of floating point precision) so use of deep networks is not beneficial, especially considering the fact that the forward pass of the network takes more time than dozen of iterations of these methods. For simple cases, the most important factor is the computation time, which is lowest for the Christensen *et al.* method [3].

In future work, we plan to train the network as a pure regularizer, as discussed in [7]. In contrast to the current approach, such a network would just add the deep regularization term during each iteration of the explicit inversion process. This approach would make the method much more universal and potentially provide even better results. Another approach could involve the use of residual links instead of the separate network branch to handle the fine details, decrease the memory consumption and the inference time.

To conclude, we propose a deep network to improve the process of inverting deformation fields in a way that minimizes the mean inverse consistency, even for deformation fields that are inherently non-invertible. The proposed method can be particularly useful for image registrations with complex deformations or methods dedicated to missing or partial data where it is known that the real deformation field is non-invertible, yet the approximate inverse can be useful.

5. ACKNOWLEDGMENTS AND COMPLIANCE WITH ETHICAL STANDARDS

This work was funded by NCN Preludium project no. UMO-2018/29/N/ST6/00143 and NCN Etiuda project no. UMO-2019/32/T/ST6/00065. Part of this study is a numerical simulation study for which no ethical approval was required. Another part was conducted retrospectively using human subject data made available in open access for the ANHIR challenge [13, 14]. The authors declare no conflict of interest.

6. REFERENCES

[1] Reaungamornrat, S. *et al.*, "MIND Demons: Symmetric Diffeomorphic Deformable Registration of MR and CT for Image-Guided Spine Surgery," *IEEE Transactions on Medical Imaging*, vol. 35, no. 11, pp. 2413–2424, 2016.

[2] Nithiananthan, S. *et al.*, "Extra-dimensional Demons: A method for incorporating missing tissue in deformable image registration," *Medical Physics*, vol. 39, no. 9, pp. 5718–5731, 2012.

[3] Christensen, G.E. and Johnson, H.J., "Consistent image registration," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 568–582, 2001.

[4] Chen, M. *et al.*, "A simple fixed-point approach to invert a deformation field," *Medical Physics*, vol. 35, no. 1, pp. 81–88, 2008.

[5] Gobbi, D.G. and Peters, T.M., "Generalized 3D non-linear transformations for medical imaging: An object-oriented implementation in VTK," *Computerized Medical Imaging and Graphics*, vol. 27, no. 4, pp. 255–265, 2003.

[6] Ardizzone, L. *et al.*, "Analyzing Inverse Problems with Invertible Neural Networks," 2018, https://arxiv.org/abs/1808.04730.

[7] Housen, L. *et al.*, "NETT: Solving Inverse Problems with Deep Neural Networks," 2018, https://arxiv.org/abs/1803.00092.

[8] Jin, K.H. *et al.*, "Deep Convolutional Neural Network for Inverse Problems in Imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[9] Mok, T. and Chung, A., "Fast Symmetric Diffeomorphic Image Registration with Convolutional Neural Networks," *IEEE CVPR*, pp. 4644–4653, 2020.

[10] Zhang, J., "Inverse-Consistent Deep Networks for Unsupervised Deformable Image Registration," *arXiv*, , no. 1809.03443, 2018.

[11] Shen, Z. and H, X. and Xu, Z. and Niethammer, M., "Networks for Joint Affine and Non-parametric Image Registration ," *IEEE CVPR*, pp. 4224–4233, 2019.

[12] He, K. *et al.*, "Identity mappings in deep residual networks," *ECCV*, vol. 9908, pp. 630–645, 2016.

[13] Borovec, J. *et al.*, "Benchmarking of Image Registration Methods for Differently Stained Histological Slides," *IEEE ICIP*, pp. 3368–3372, 2018.

[14] J. Borovec et al., "ANHIR: Automatic Non-rigid Histological Image Registration Challenge," *IEEE Transactions on Medical Imaging*, vol. 39, no. 10, pp. 3042–3052, 2020.