Networks for Joint Affine and Non-parametric Image Registration

Zhengyang Shen
UNC Chapel Hill
zyshen@cs.unc.edu

Xu Han
UNC Chapel Hill
xhs400@cs.unc.edu

Zhenlin Xu
UNC Chapel Hill
zhenlinx@cs.unc.edu

Marc Niethammer
UNC Chapel Hill
mn@cs.unc.edu

A. Supplementary material

This supplementary material provides additional details illustrating the proposed approach. Specifically, Sec. A.1 describes how the affine training is regularized in an epoch-dependent way. Sec. A.2 shows registration performance for different numbers of steps for the affine registration-part of the network. Sec. A.3 details the structure of the momentum generation network. Lastly, Sec. A.4 shows additional registration examples.

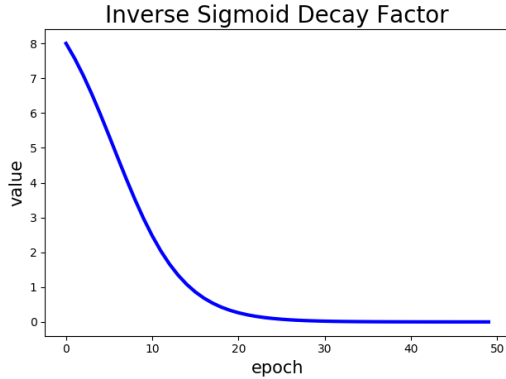### A.1. Affine regularization factor

Figure 1. Graph of the affine regularization factor. Its value decays to zero over the epochs.

To help with convergence of the affine registration network, we use an epoch-dependent regularization factor, which discourages large transformations at the start of the training. Specifically, we define this epoch-dependent regularization factor as

$$\lambda_{ar} := \frac{C_{ar}K_{ar}}{K_{ar} + e^{n/K_{ar}}}, \quad (1)$$

where C_{ar} is a constant, K_{ar} controls the decay rate, and n is the epoch count. Fig. 1 shows the value of λ_{ar} plotted over the epochs. As the value decays to zero with the epochs, its influence on the training becomes negligible. For both longitudinal and cross-subject experiments, K_{ar} is set to 4 and C_{ar} is set to 10.
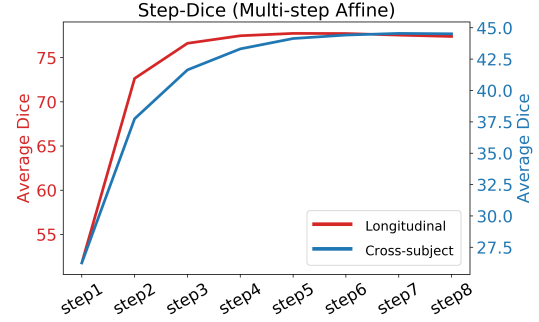

Figure 2. Multi-step Affine registration results over iteration steps. The affine network is trained using three steps for longitudinal registration (red) and five steps for cross-subject registration (blue). Performance increases with steps and finally saturates.

A.2. Dice over steps in Multi-step Affine Network

The main manuscript shows the average Dice scores over the number of test iteration steps for the vSVF registration component. For completeness, Fig. 2 shows the average Dice scores over the number of steps for the affine network. The model is trained using a three-step affine network for longitudinal registrations and using five steps for cross-subject registration. Similar to the vSVF registration, it can be observed that model performance improves with large number of steps and saturates at a high performance level.

A.3. Structure of Momentum Generation Network

As the network structure itself is not the main contribution of our work, we do not describe it in detail in the main manuscript. For completeness, we describe the architecture here. Fig. 3 shows the structure of the Momentum Generation Network. It takes a pair of images as the input and outputs a low-resolution initial momentum. We use a four-level U-net [4, 2] with residual links, but remove the last decoder level to output the low-resolution momentum. As the momentum can be positive or negative, no activation function (*e.g.* ReLu [3] or leakyRelu [1]) is used after the last two convolutional layers, which output the momentum.
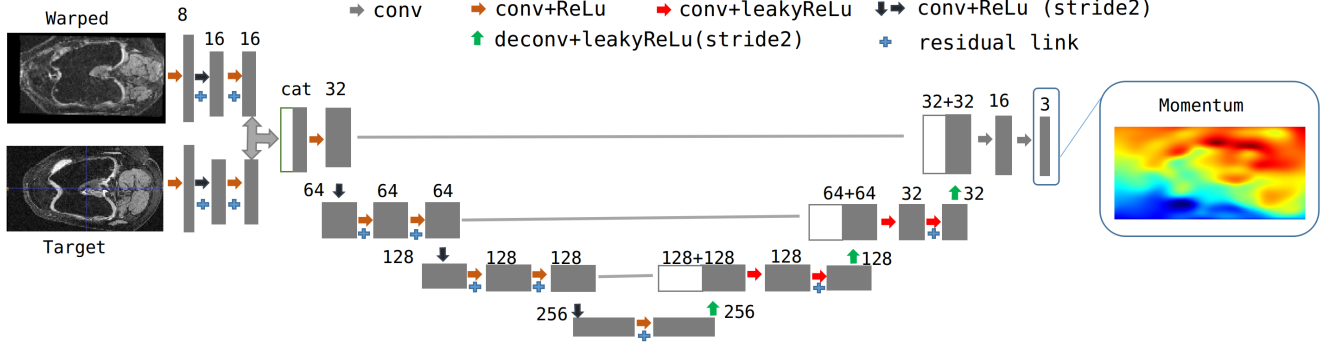
Figure 3. Illustration of the structure of Momentum Generation Network. It follows the structure of the U-net but the last level decoder is removed.

A.4. Visualization

To provide more insight into the registration behavior of our network, we visualize results illustrating deformation folds, results for different steps in the multi-step approach, and additional examples. Specifically, we show the following:

- *Folds*: To better visualize the folds produced by the multi-step vSVF, we report the registration results, shown in Fig. 4, from the six-step vSVF. These folds mostly occur at regions of anatomical inconsistency or at the image boundary where map interpolation artifacts may influence the solution. In these regions, very large momentum values may be predicted which can result in folds due to discretization artifacts when integrating the advection equation.

- *Multi-step in vSVF registration*: Fig. 5 shows the registration results over the steps of the vSVF. Although folds may result from the multi-step strategy in some very large deformation cases, the transformation maps are largely well regularized. We observe that the registration results improve over the steps.

- *More AVSM examples*: Fig. 6 shows additional AVSM registration results. It can be observed that AVSM achieves good registration results with smooth transformation maps for cases with large and small deformations.

References

[1] A. L. Maas, A. Y. Hannun, and A. Y. Ng. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3, 2013.

[2] F. Milletari, N. Navab, and S.-A. Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *3D Vision (3DV), 2016 Fourth International Conference on*, pages 565–571. IEEE, 2016.

[3] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814, 2010.

[4] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241. Springer, 2015.
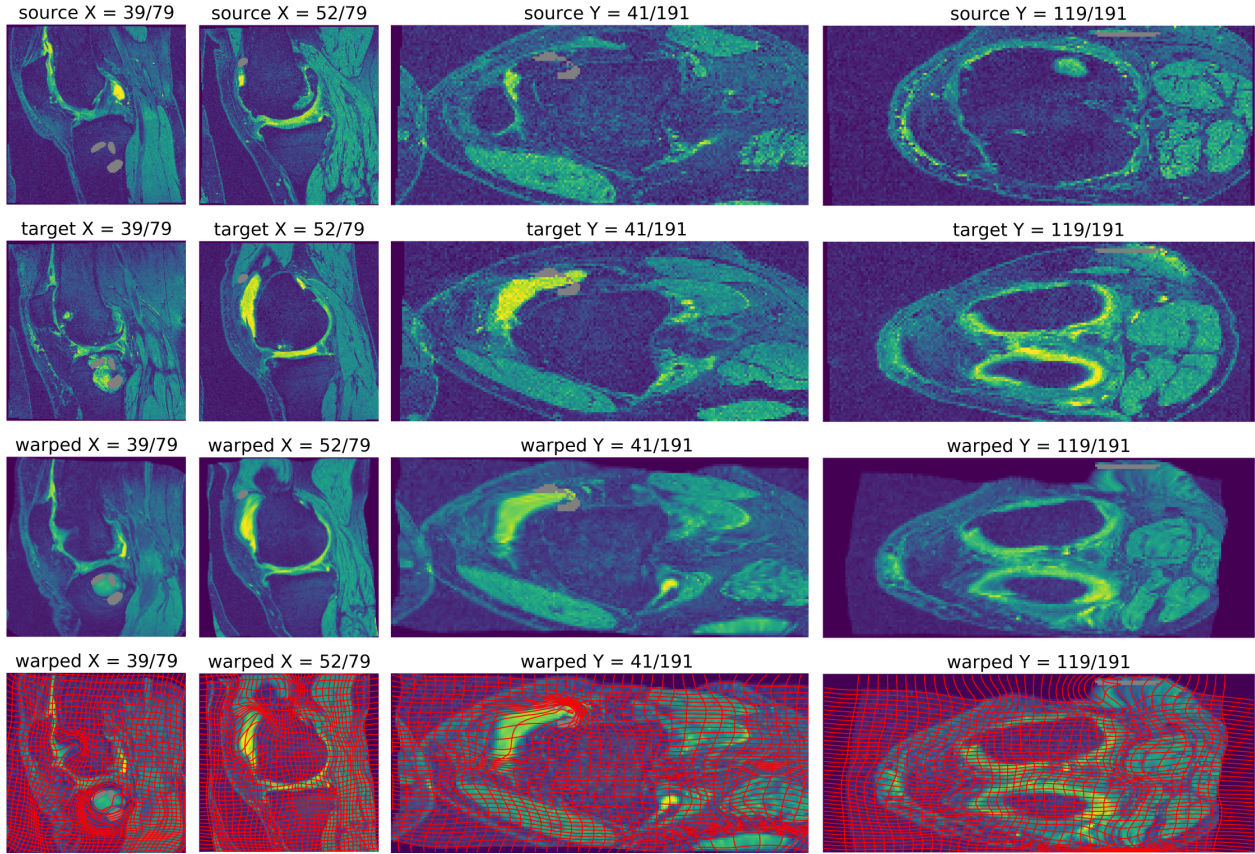
Figure 4. Examples of folds produced by a six-step vSVF (trained using a two-step vSVF). Each column refers to an example registration case. From top to bottom source, target, warped image by AVSM and warped image with deformation grid (visualizing Φ^{-1}) are shown. Folds are shown in *gray*. From left to right, the first three columns refer to cases with anatomical inconsistency and the last column refers to a case where the folds occur at the boundary.
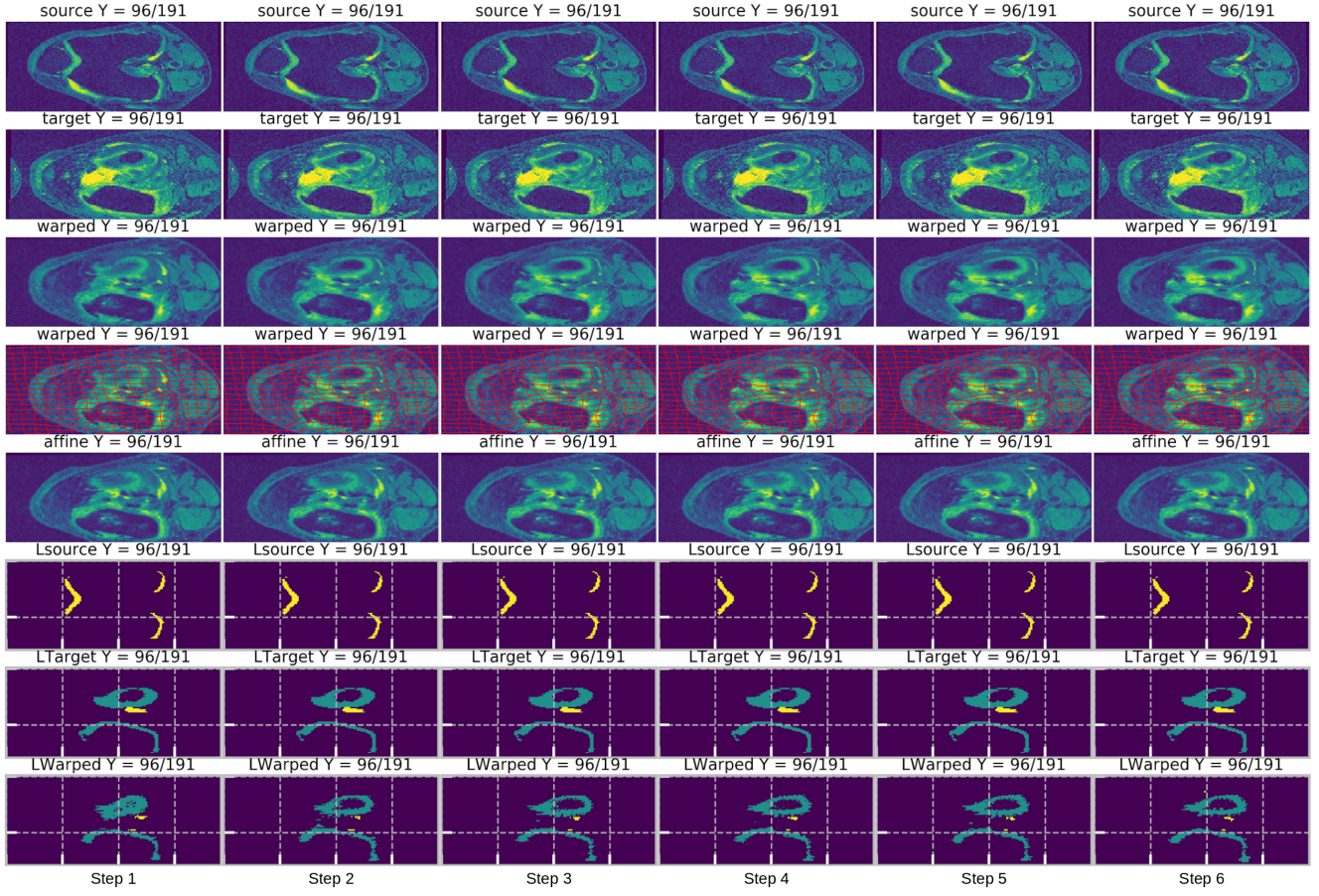
Figure 5. Illustration of the results of *one registration case (with six steps)* by AVSM (trained using a two-step vSVF). From left to right, each column shows results for different steps. The first five rows refer to source, target, warped image by AVSM, warped image with deformation grid (visualizing Φ^{-1}) and warped image by the multi-step affine network respectively. The last three rows show the source label, target label and warped label for the AVSM result. The transformation map gets refined over the six steps and the registration result improves as indicated by a better correspondence between the target label and the warped label images (last two rows).
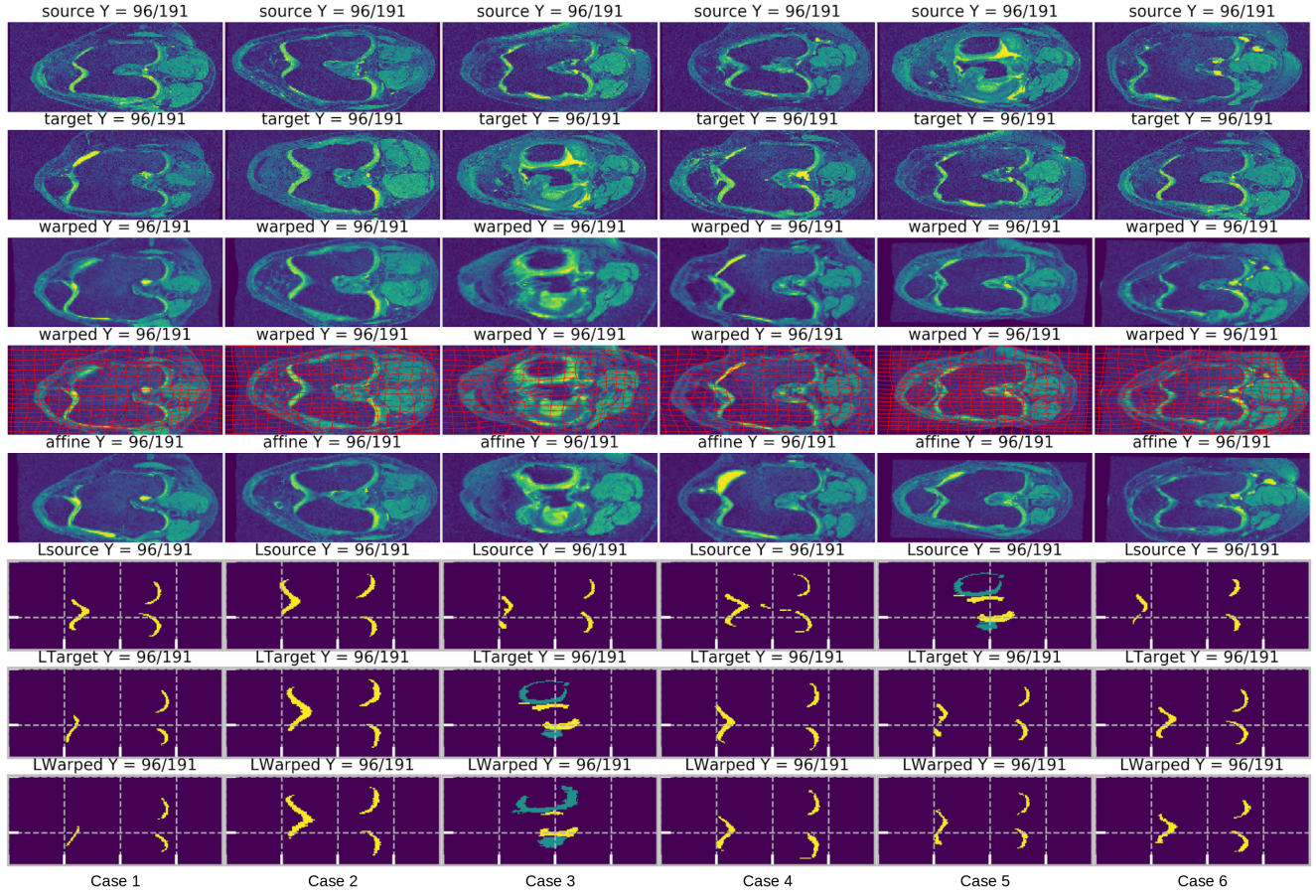
Figure 6. Illustration of results of *six registration cases* by AVSM. Each column refers to an example registration case. The first five rows refer to source, target, warped image by AVSM, warped image with deformation grid (visualizing Φ^{-1}) and warped image by the multi-step affine network respectively. The last three rows show the source label, target label and warped label by AVSM. There is high similarity between the warped and the target images and the deformations are smooth, illustrating the good registration performance of our proposed AVSM approach.