GSMorph: Gradient Surgery for cine-MRI Cardiac Deformable Registration

Haoran Dou[1], Ning Bi[1], Luyi Han[3,4], Yuhao Huang[5], Ritse Mann[3,4], Xin Yang[5,6], Dong Ni[5], Nishant Ravikumar[1], Alejandro F. Frangi[1,2,7,8], and Yunzhi Huang[9]

1. University of Leeds, UK 2. University of Manchester, UK 3. Radboud University Medical Centre, The Netherlands
4. Netherlands Cancer Institute, The Netherlands 5. Shenzhen University, China 6. Shenzhen RayShape Medical Technology Co., Ltd, China
7. Katholieke Universiteit Leuven, Belgium 8. Alan Turing Institute, UK 9. Nanjing University of Information Science and Technology, China

Haoran Dou

Background

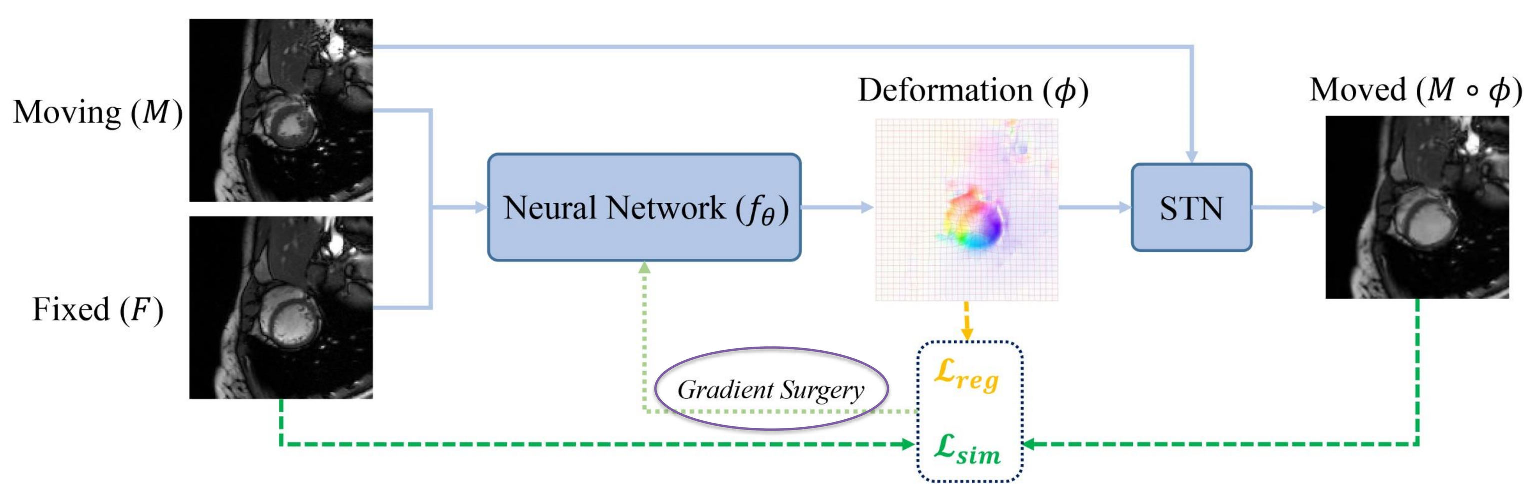
Deep-learning based registration (DLR) formulates registration as a network learning process minimizing a composite objective function comprising one similarity loss to penalize the difference in the appearance of the image pair, and a regularization term to ensure the smoothness of deformation field (DF). Performing hyperparameter tuning to balance the registration accuracy and smoothness of DF is typically labor-intensive, time-consuming, and *ad-hoc*; searching for the optimal parameter setting requires extensive ablation studies and hence training tens of models and establishing a reasonable parameter search space. Therefore, **alleviating, even circumventing, hyper parameter search to accelerate development and deployment of DLR models** remains challenging.

Highlights

- We propose GSMorph, a **gradient-surgery-based DLR model,** to circumvent tuning the hyperparameter in composite loss function with a gradient-level reformulation to reach the trade-off between registration accuracy and smoothness of the deformation field.
- We propose a **layer-wise GS** to group by the parameters for optimization to ensure the flexibility and robustness of the optimization process.
- Our method is **model-agnostic** and can be integrated into any DLR network without extra parameters or losing inference speed.

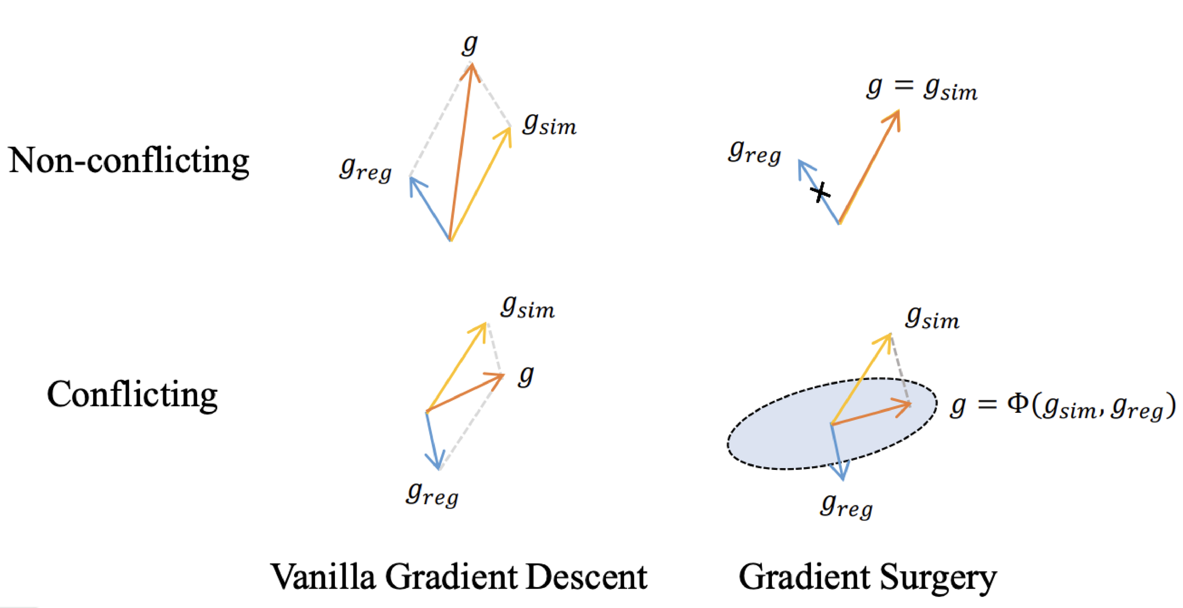
Methods

①Gradient Surgery-based Parameters Optimization



Before $\theta := \theta - \nabla_\theta(\mathcal{L}_{sim}(\theta; F, M \circ \phi) + \mathcal{L}_{smooth}(\theta; \phi))$

After $\theta := \theta - \alpha\Phi(\nabla_\theta(\mathcal{L}_{sim}(\theta; F, M \circ \phi))$ $\Phi(\cdot)$: *Gradient Surgery*



The left Figure visualizes the vanilla gradient descent and gradient surgery for non-conflicting and conflicting gradients.

The conflicting relationship between two controversial constraints in the loss function can be geometrically projected as orthogonal vectors. Depending on the orthogonal relationship, merely updating the gradients of the similarity loss would automatically associate with the updates of the regularization term. In this way, we avoid tuning the hyperparameter that balances the tradeoff between the similarity loss and the regularization loss to optimize θ. The specific strategy is shown in the Algorithm 1

②Layer-wise Gradient Surgery

Algorithm 1 Gradient surgery

Require: Parameters θ_i in ith layer of the network; Number of layers in the network N.

$g_{sim} \leftarrow \nabla_\theta \mathcal{L}_{sim}$
$g_{reg} \leftarrow \nabla_\theta \mathcal{L}_{reg}$
for $i = 1 \rightarrow N$ **do**
  if $\langle g^i_{sim}, g^i_{reg} \rangle > 0$ **then**
    $g_i = g^i_{sim}$
  else
    $g_i = g^i_{sim} - \frac{\langle g^i_{sim}, g^i_{reg} \rangle}{\|g^i_{reg}\|^2} g^i_{reg}$
  end if
  $\Delta\theta_i = g_i$
end for
Update θ

Compared with existing studies performing the GS in terms of independent parameters or the entire network. we introduce a layer-wise GS to ensure its stability and flexibility. The gradient surgery is performed for each layer in the neural network during optimization.

Experimental Results

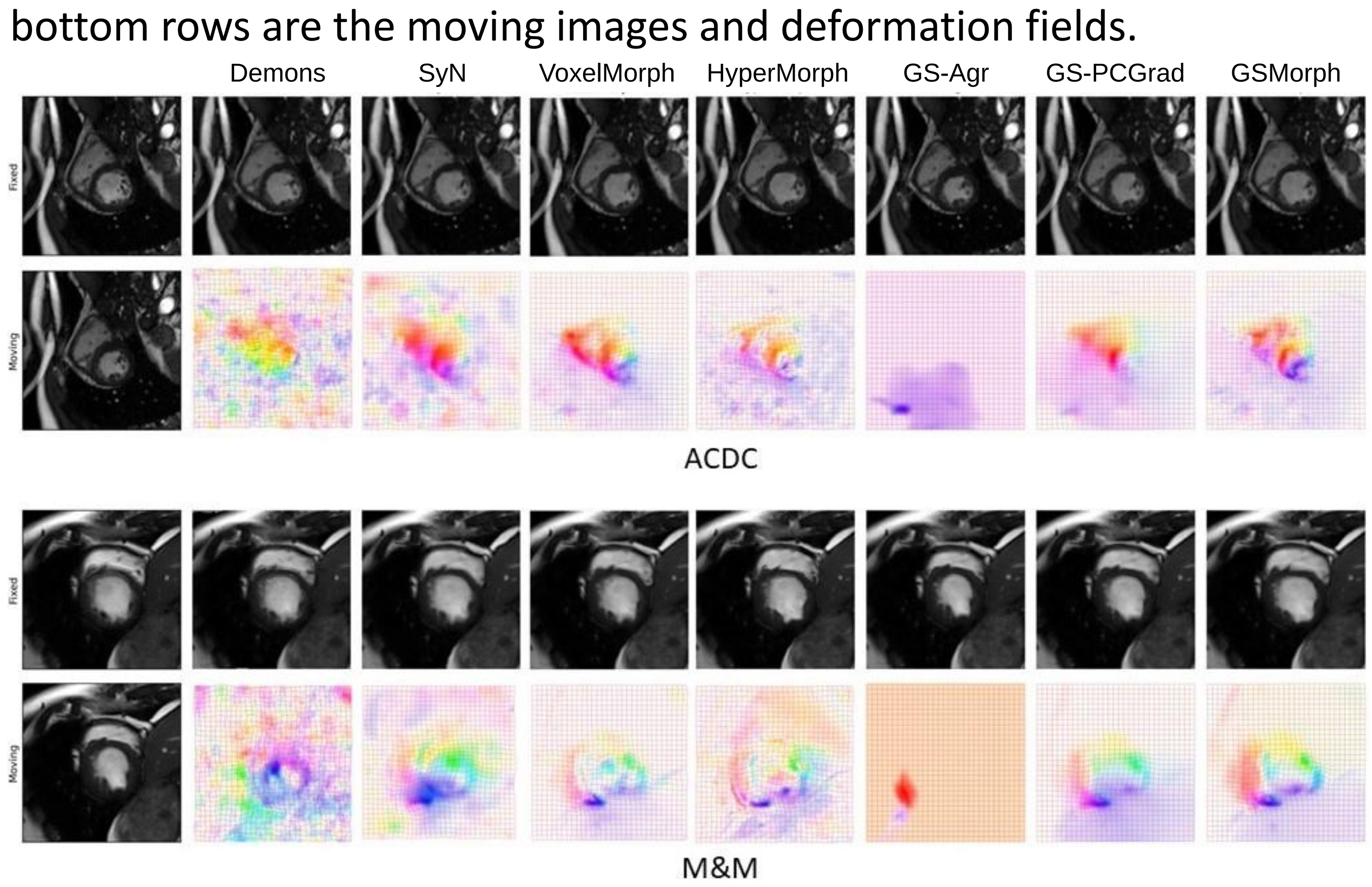
Dataset

ACDC: 100 subjects; M&M: 219 subjects

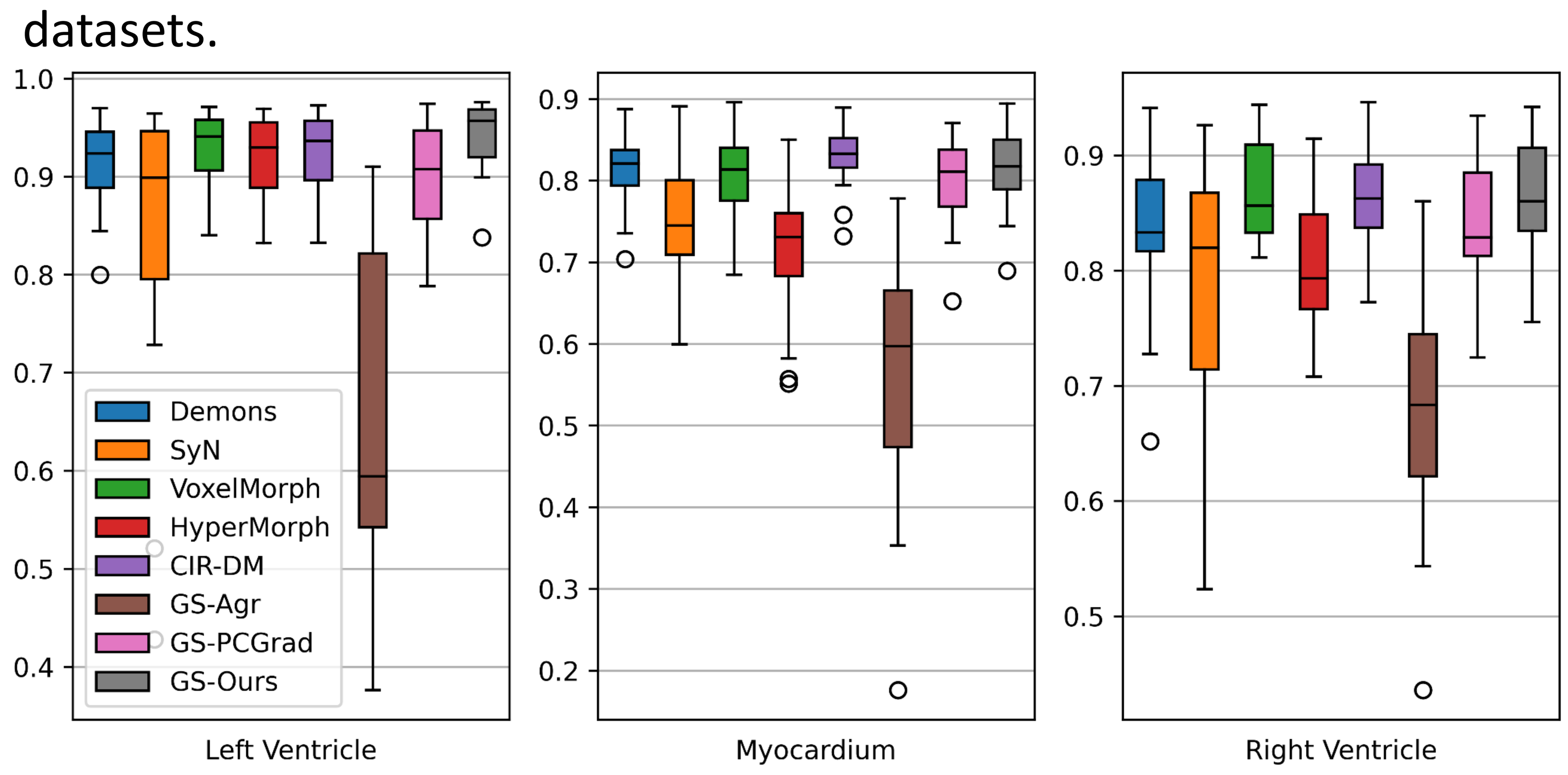
Moving Image: End Systole (ES); Fixed Image: End Diastole (ED)

Results

Visual comparison of the registration results of the investigated methods for two test cases in ACDC and M&M datasets. The top rows are the fixed images and moved images from different methods; the bottom rows are the moving images and deformation fields.



Boxplot of the registration results (Dice) of the competitive methods in terms of Left Ventricle, Myocardium, and Right Ventricle in ACDC datasets.



Conclusion

This work presents the first gradient-surgery-based framework for alleviating the hyperparameter tunning in medical image registration. We believe our method will facilitate the development and deployment of registration models and alleviate the influence of hyperparameters on performance.