Multi-scale Neural ODEs for 3D Medical Image Registration

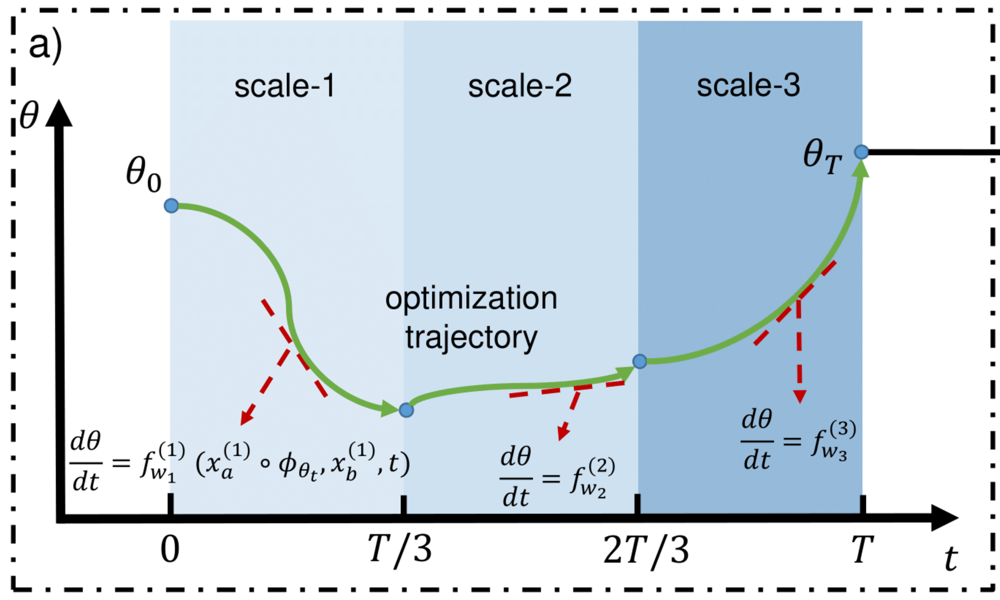
Junshen Xu, Eric Z. Chen, Xiao Chen, Terrence Chen, and Shanhui Sun
Junshen@mit.edu

Introduction

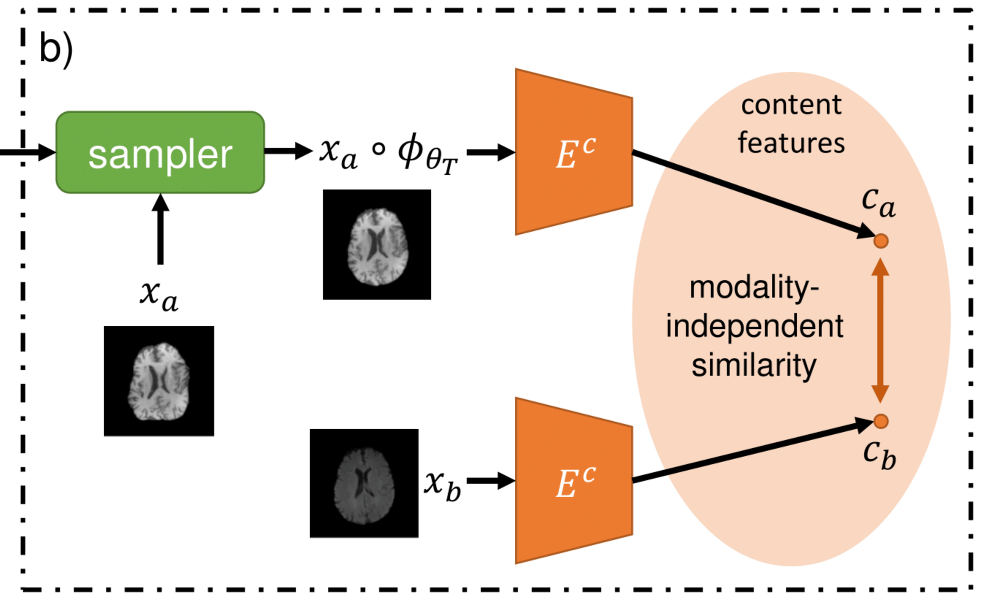
- We proposed a new direction of modeling image registration optimizer as a continuous optimization dynamics via neural ODEs.
- We introduced multi-scale architecture to neural ODEs to reduce searching space by performing registration iteratively on different scales.
- Our proposed method is a general learn-to-learn image registration framework and is not limited to specific transformations.
- Our framework can handle multiple contrasts with a single trained network attribute to proposed contrast-independent similarity metric for n(≥ 2) modalities.

Methods

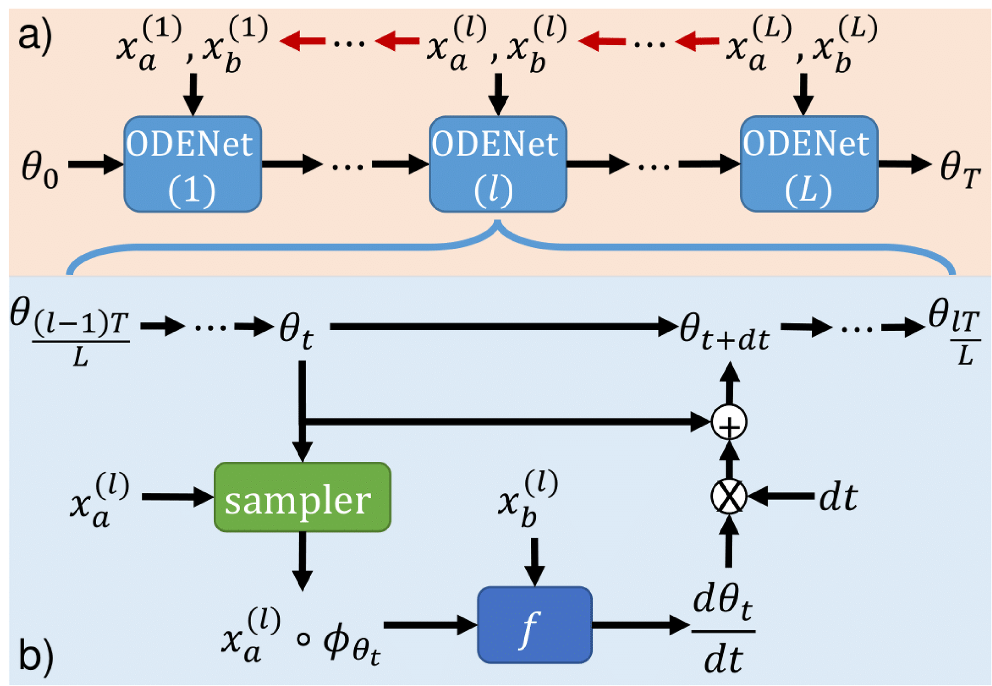
- Image registration is an optimization problem:
$$\operatorname*{argmin}_{\theta} L(x_a \circ \phi_\theta, x_b) + R(\theta)$$
- Solve with gradient descent based algorithms:
$$\theta_{t+1} = \theta_t - \eta_t \frac{\partial(L+R)}{\partial \theta_t} \triangleq \theta_t + f(\theta_t, t)$$
- Consider the optimization process as a continuous flow in time:
$$\frac{d\theta_t}{dt} = f(\theta_t, t), t \in [0, T]$$
- Given the initial parameter θ_0, the final parameter θ_T is the solution to this ODE initial value problem:
$$\theta_T = \theta_0 + \int_0^T f(\theta_t, t) dt$$
- We choose the following form of ODE:
$$\frac{d\theta_t}{dt} = f_w(x_a \circ \phi_{\theta_t}, x_b, t), t \in [0, T]$$
- Multi-scale ODE
$$\frac{d\theta_t}{dt} = f^{(l)}_{w_l}\left(x_a^{(l)} \circ \phi_{\theta_t}, x_b^{(l)}, t\right), t \in [\frac{l-1}{L}T, \frac{l}{L}T]$$
- Multi-modal/contrast registration
  - M different modality/contrast groups, $\{X_i\}_{i=1}^M$
  - $x_a \in X_a, x_b \in X_b$
- Modal-independent loss
  - Multimodal image translation with content and style feature disentanglement
  - Use the retrained content encoder E^c to extract modal-independent features
$$L_{\text{sim}} = \mathrm{E}_{x_a, x_b, a, b} \left\| E^c(x_a \circ \phi_{\theta_T}) - E^c(x_b) \right\|_2^2$$



The image registration optimization is modeled as a neural ODE



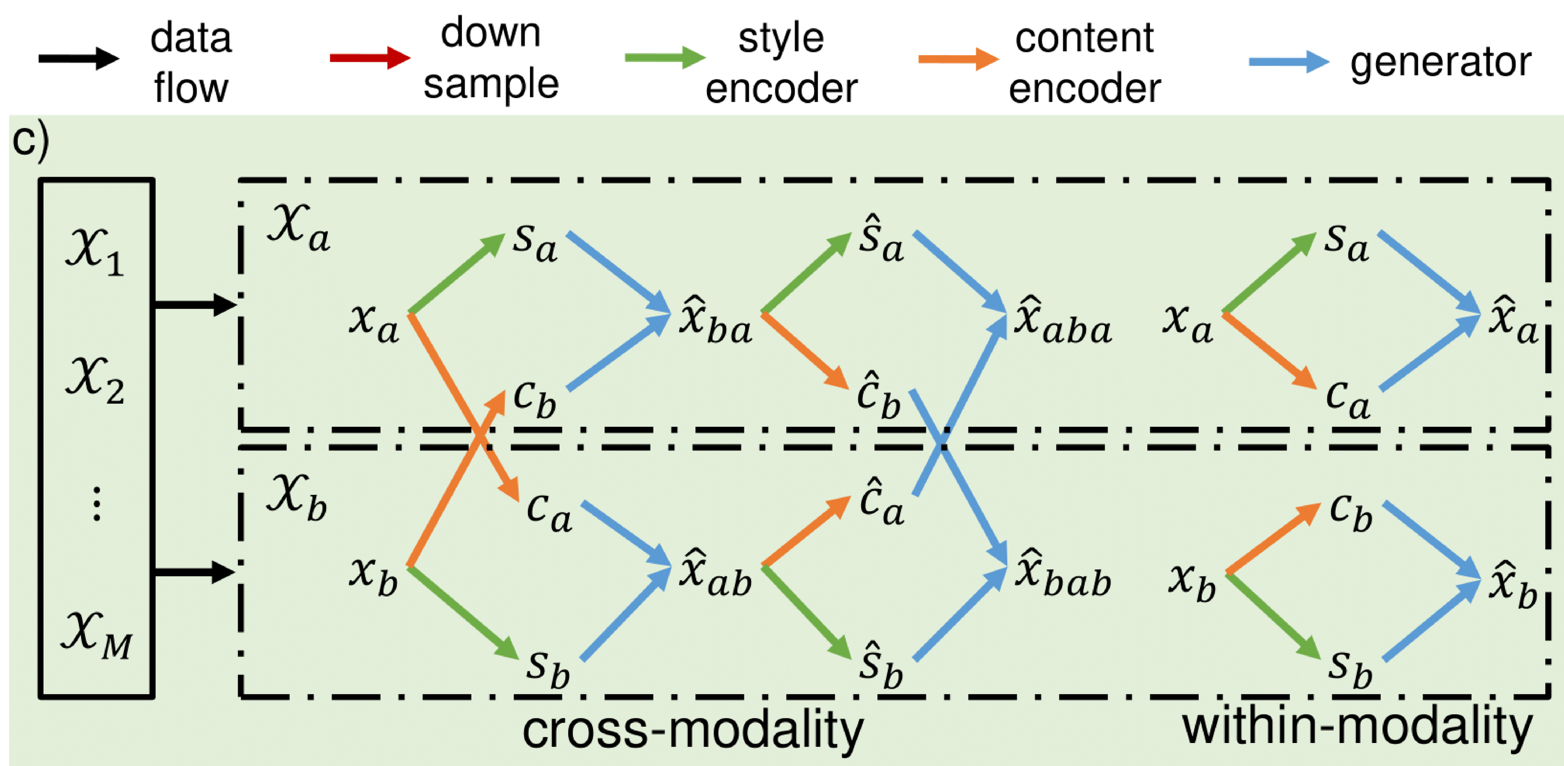
The modality-independent similarity metric is realized via a pretrained encoder



Multi-scale neural ODE

Choices of network f for different transformations

- **Rigid/Affine:** convolution networks with fully connected layers
- **B-spline:** convolution networks with downsampling
- **Dense motion field:** UNet based networks



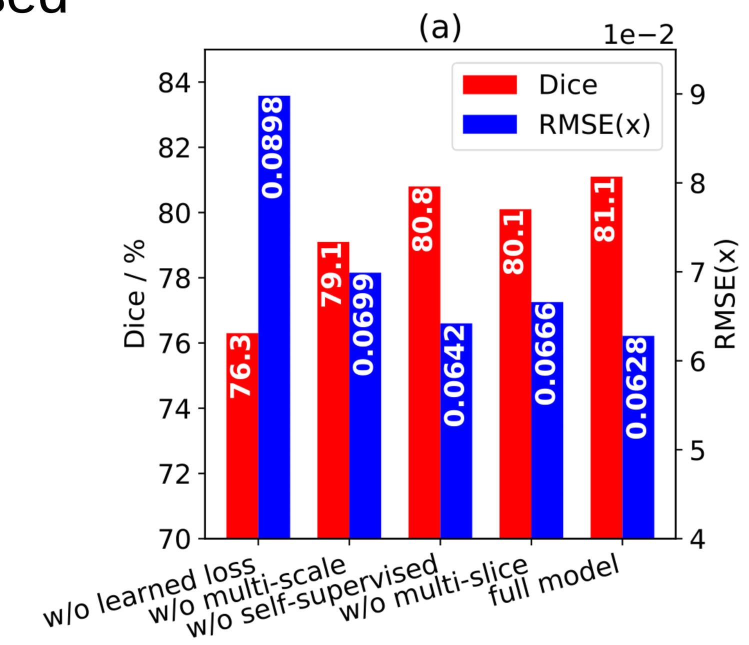
overview of the utilized image translation towards metric learning

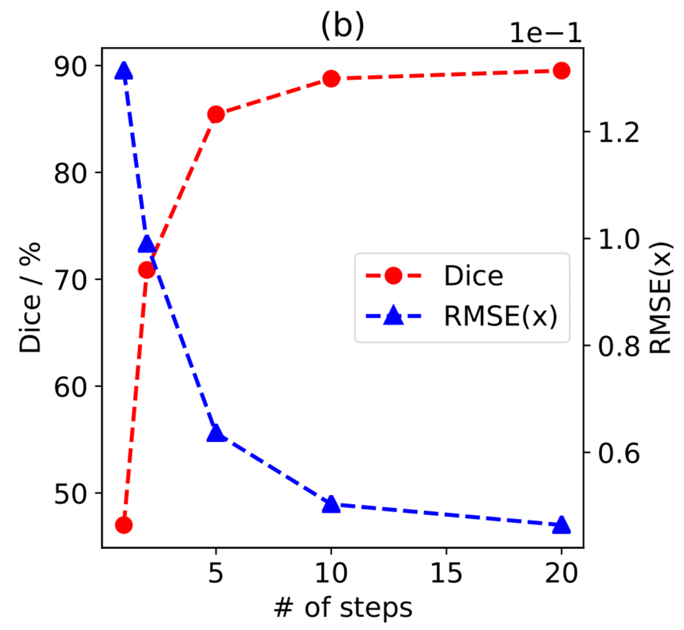
Experiments

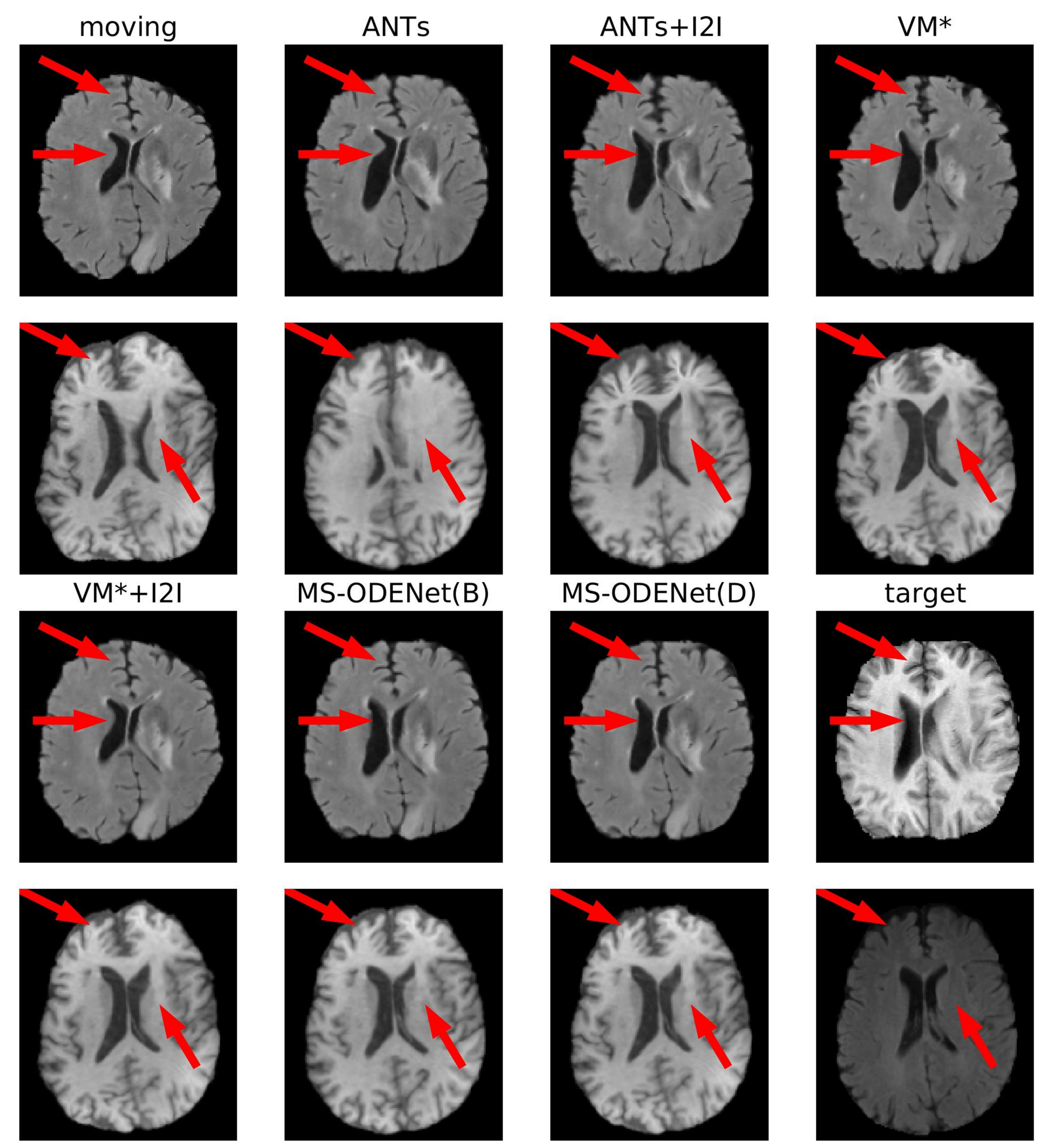
- **Dataset: BraTS 2020**
  - multi-modal 3D brain MRI
  - Four distinctive contrasts: T1, T2, T2-FLAIR, and T1Gd
  - 494 subjects with glioblastomas
- **Neural ODE solver**
  - Euler's method
  - Adaptive Heun's method
- **Transformation**
  - Rigid
  - Deformable
  - rigid + deformable

Results

Transformation	Methods	Dice/%	RMSE(x)	RMSE(ϕ)/mm	Time/s
Rigid	ANTs	63.0	8.34	7.28	17.17
	ANTs+I2I	60.6	8.83	7.54	30.69
	MS-ODENet(R)	**90.6**	**3.89**	**3.57**	**0.55**
Deformable	ANTs	81.9	6.31	1.21	55.35
	ANTs+I2I	81.1	6.34	1.06	69.47
	VM	79.4	8.81	1.61	**0.24**
	VM+I2I	80.1	8.52	1.26	0.34
	MS-ODENet(D)	81.6	6.63	1.11	1.13
	MS-ODENet(B)	**83.0**	**6.17**	**0.99**	0.31
Rigid + Deformable	ANTs	73.6	6.79	2.99	70.87
	ANTs+I2I	71.1	6.97	2.83	87.25
	RCN	64.9	8.74	5.48	2.49
	VM*	78.1	8.62	2.92	0.60
	VM*+I2I	78.4	7.07	2.04	1.01
	MS-ODENet(R+D)	79.6	6.70	1.82	1.39
	MS-ODENet(R+B)	**81.1**	**6.28**	**1.52**	**0.56**







Conclusions

- We present a new framework for 3D multi-modal image registration.
- Experiment results show that our proposed framework is superior to other compared methods.
- For future work, we will extend our framework to other types of medical registration such as 3D-2D image registration.