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: argmin f (x_a \circ \phi_{\theta_0}, x_b) + R(\theta)
- Solve with gradient descent based algorithms: \theta_{t+1} = \theta_t - \eta \frac{\partial L}{\partial \theta_t} = \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 \theta_0, the final parameter \theta_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_\omega(x_a \circ \phi_{\theta_t} x_b, t), t \in [0, T]
- Multi-scale ODE
  \frac{d\theta_t}{dt} = \sum_{l=1}^L \frac{1}{L} f_\omega(x_a^{(l)} \circ \phi_{\theta_t} x_b^{(l)}, t), t \in [\frac{l-1}{L}, \frac{l}{L}]
- 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_{sim} = E_{x_{a,b} \circ \phi_{\theta_T}} \| E_c(x_a \circ \phi_{\theta_T}) - E_c(x_b) \|^2_z

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

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Methods</th>
<th>Dice%</th>
<th>RMSE(x)</th>
<th>RMSE(\phi/\mu m)</th>
<th>Time/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rigid</td>
<td>ANTs</td>
<td>63.0%</td>
<td>8.34</td>
<td>7.28</td>
<td>17.17</td>
</tr>
<tr>
<td></td>
<td>ANTs+I2I</td>
<td>60.6%</td>
<td>8.83</td>
<td>7.54</td>
<td>30.69</td>
</tr>
<tr>
<td></td>
<td>MS-ODENet(R)</td>
<td>90.6%</td>
<td>3.89</td>
<td>3.57</td>
<td>0.55</td>
</tr>
<tr>
<td>Deformable</td>
<td>ANTs</td>
<td>81.9%</td>
<td>6.31</td>
<td>1.21</td>
<td>55.35</td>
</tr>
<tr>
<td></td>
<td>ANTs+I2I</td>
<td>81.1%</td>
<td>6.34</td>
<td>1.06</td>
<td>69.47</td>
</tr>
<tr>
<td></td>
<td>VM</td>
<td>79.4%</td>
<td>8.81</td>
<td>1.61</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>VM+I2I</td>
<td>80.1%</td>
<td>8.52</td>
<td>1.26</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>MS-ODENet(D)</td>
<td>81.6%</td>
<td>6.63</td>
<td>1.11</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>MS-ODENet(B)</td>
<td>83.0%</td>
<td>6.17</td>
<td>0.99</td>
<td>0.31</td>
</tr>
</tbody>
</table>

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.