Quicksilver: fast deep learning registration

Yang et al, Neuroimage 2017

Jan Kybic

2020

Key features

- Deep learning-based
- Prediction/correction network
- Fast (11s on 1GPU for a 3D volume)
- Diffeomorphic transformation
- Large deformations (large deformation diffeomorphic metric mapping - LDDMM)
- Patch-based, patch pruning
- Uncertainty quantification
- Multimodal registration

Formulation

$$E(\Phi) = \mathrm{Reg}[\Phi] + \frac{1}{\sigma^2}\mathrm{Sim}[I_0 \circ \Phi^{-1}, I_1]$$

LDDMM is a non-parametric registration method which represents the transformation via spatio-temporal velocity fields. In particular, the sought-for mapping, Φ, is obtained via an integration of a spatio-temporal velocity field $v(x,t)$ for unit time, where t indicates time and $t \in [0,1]$, such that $\Phi_t(x,t) = v(\Phi(x,t),t)$ and the sought-for mapping is $\Phi(x,1)$. To single-out desirable velocity-fields, non-

$$E(v) = \int_0^1 \|v\|_L^2 \, dt + \frac{1}{\sigma^2}\|M \circ \Phi^{-1}(1) - T\|^2,$$

$$\text{s.t.} \quad \Phi_t(x,t) = v(\Phi(x,t),t), \;\; \Phi(x,0) = \mathrm{id}$$

Differential formulation $\Phi_t^{-1} + D\Phi^{-1}v = 0$

Classical solution

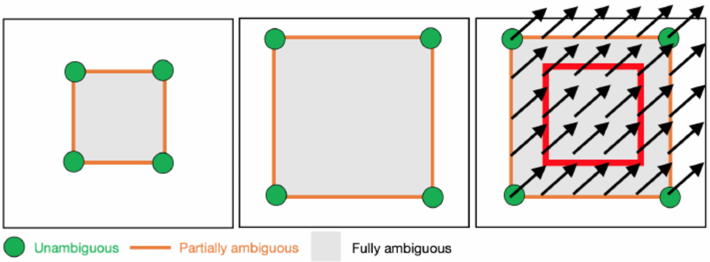
- forward transformation - follow a particle in v. Ensures diffeomorphy
- optimization
  - current mismatch
  - solve (adjoint) system backward
  - gradient of the velocity field at all t
  - update v

Shooting formulation

- find the shortest path (geodesics) between images
- geodesic parameterized by initial $\Phi^{-1}$and momentum m= Lv
- m supported mainly on image edges, $m(x,t) = \lambda(x,t)\nabla I(x,t)$
- v is a smoothed momentum, $v = L^{-1}m$

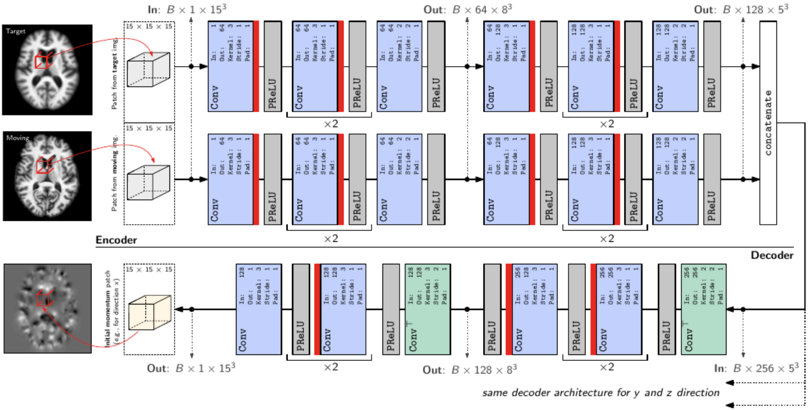
Method

- predict m patch-by-patch
- train network to predict m
- training data - m found by numerical optimization
- m well predicted from patches, does not have to be smooth, zero in homogeneous regions



- large stride, drop background patches

Network structure



encoder/decoder, $l_1$loss function on m (not E),
3 decoders (easier to train)

Probabilistic network

- Instead of $\boldsymbol{y} = f(\boldsymbol{x})$, predict $p(\boldsymbol{y}|\boldsymbol{x}.\boldsymbol{X}, \boldsymbol{Y})$for training data $\boldsymbol{X}, \boldsymbol{Y}$
- variational inference for network weights $\boldsymbol{W}$,minimize KL divergence of $q(\boldsymbol{W})$ and $p(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y})$

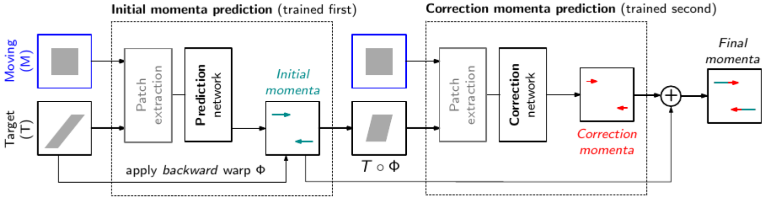
$$q(\mathbf{W}_i) = \mathbf{M}_i \cdot \mathrm{diag}([z_{i,j}]_{j=1}^{K_i}), \quad z_{i,j} \sim \mathrm{Bernoulli}(d) \ ,$$

- $\rightarrow$dropout with probability 0.2

$$p(\mathbf{y}'|\mathbf{x}', \mathbf{X}, \mathbf{Y}) \approx \frac{1}{T}\sum_{t=1}^{T} \hat{f}(\mathbf{x}', \hat{\mathbf{w}})$$

- result=mean, variance $\rightarrow$uncertainty estimate

Prediction/correction

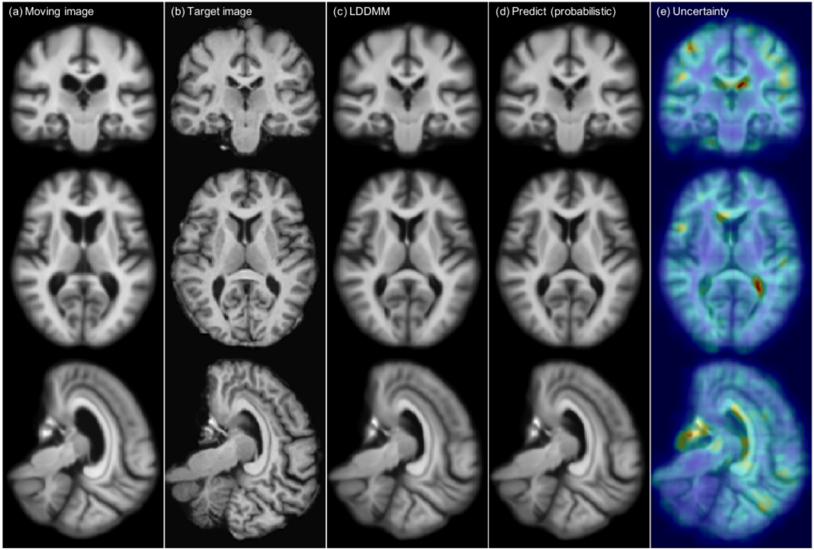


- trained sequentially
- $M$and $T \circ \Phi$ are in the same coordinate space, can be added

Datasets

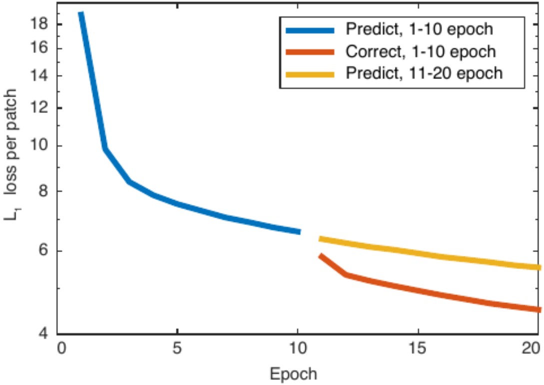
- T1, T2 MR images
- training m obtain from T1 images $\rightarrow$learn also multimodal T1-T2 registration

Atlas-to-image example



blue - low uncertainty

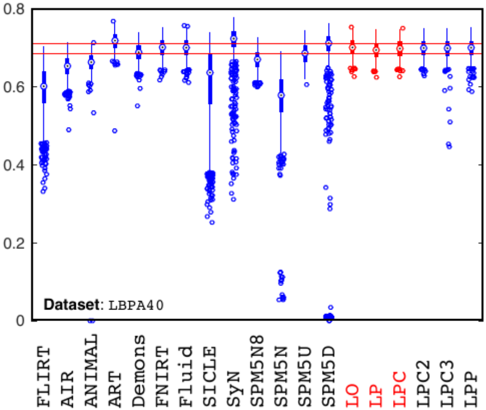
Prediction/correction experiments



Quantitative results

	Deformation Error for each voxel [mm]							$\mathbf{det}J > \mathbf{0}$
Data percentile for all voxels	0.3%	5%	25%	50%	75%	95%	99.7%	
`Affine`	0.0613	0.2520	0.6896	1.1911	1.8743	3.1413	5.3661	N/A
`D`, velocity, stride 5	0.0237	0.0709	0.1601	0.2626	0.4117	0.7336	1.5166	100%
`D`, velocity, stride 14	0.0254	0.075	0.1675	0.2703	0.415	0.743	1.5598	100%
`D`, deformation, stride 5	0.0223	0.0665	0.1549	0.2614	0.4119	0.7388	1.5845	56%
`D`, deformation, stride 14	0.0242	0.0721	0.1671	0.2772	0.4337	0.7932	1.6805	0%
`P`, **momentum, stride 14, 50 samples**	0.0166	0.0479	0.1054	0.1678	0.2546	0.4537	1.1049	100%
`D`, **momentum, stride 5**	0.0129	0.0376	0.0884	0.1534	0.2506	0.4716	1.1095	100%
`D`, **momentum, stride 14**	0.013	0.0372	0.0834	0.1359	0.2112	0.3902	0.9433	100%
`D`, **momentum, stride 14, 40 epochs**	0.0119	0.0351	0.0793	0.1309	0.2070	0.3924	0.9542	100%
`D`, **momentum, stride 14 + correction**	**0.0104**	**0.0309**	**0.0704**	**0.1167**	**0.185**	**0.3478**	**0.841**	100%

Target overlap



Multimodal registration

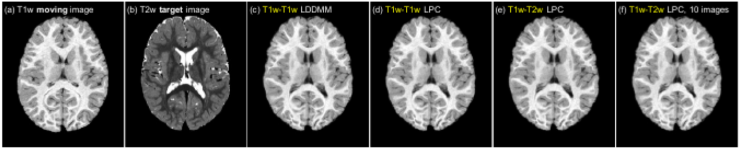


Figure 10: Example test case for *multi-modal image-to-image* tests. (a): T1w moving image; (b): T2w target image; (c): T1w-T1w LDDMM optimization (`LO`) result; (d)-(f): deformation prediction+correction (`LPC`) result using (d) T1w-T1w data; (e) T1w-T2w data; (f) T1w-T2w data using only 10 images as training data.