Amsterdam UMC
Universitair Medische Centra

Unsupervised deep learning image registration: Beyond the cranial vault

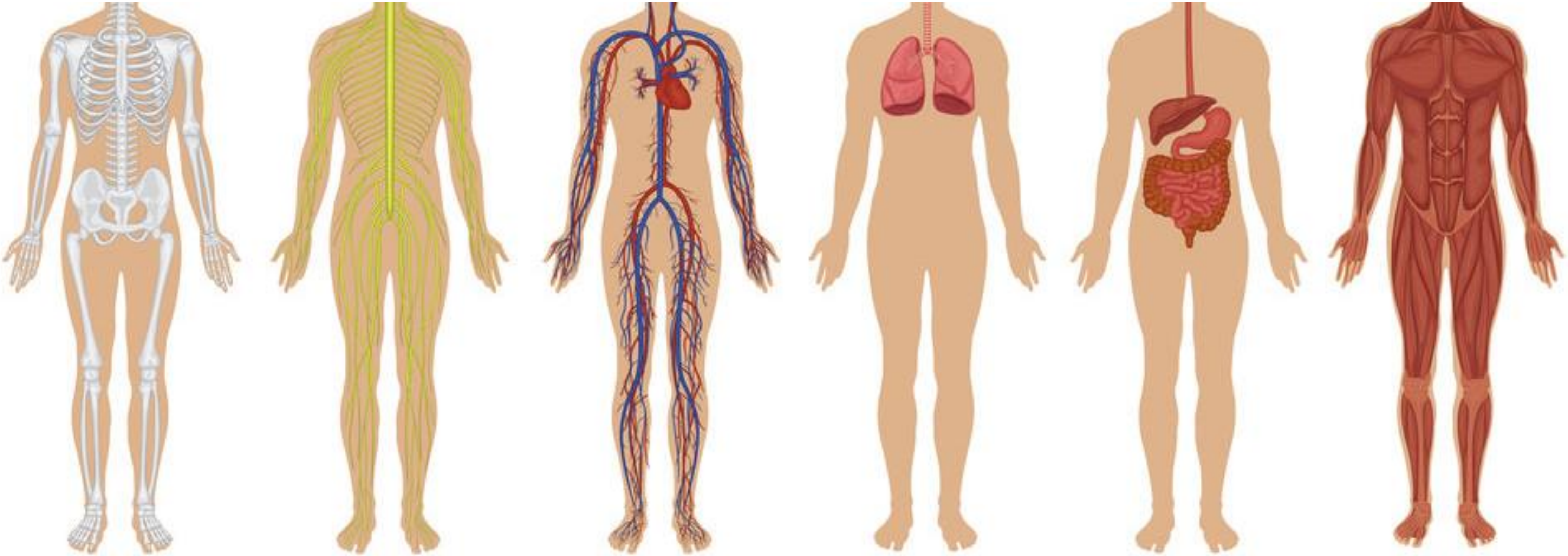
Bob de Vos

Postdoc Amsterdam UMC

Learn2Reg



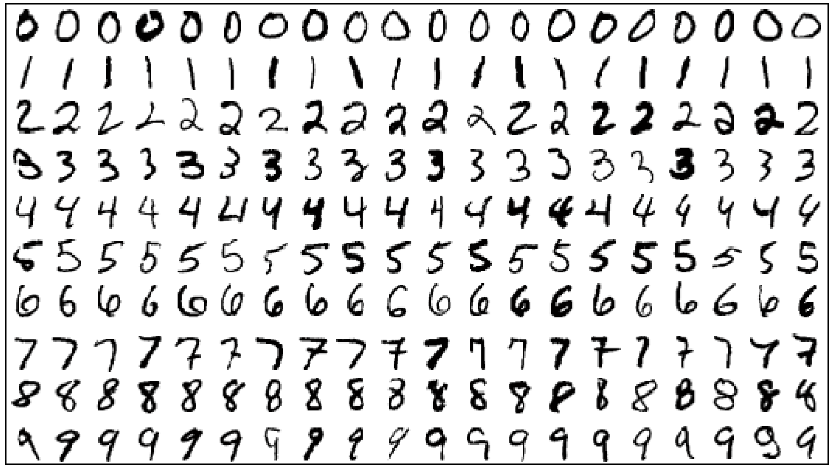


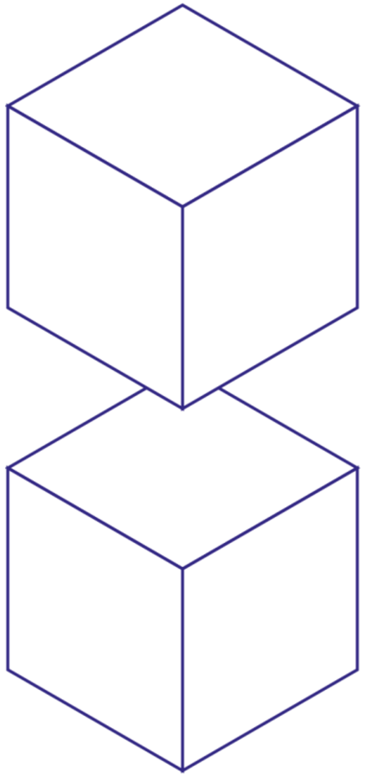






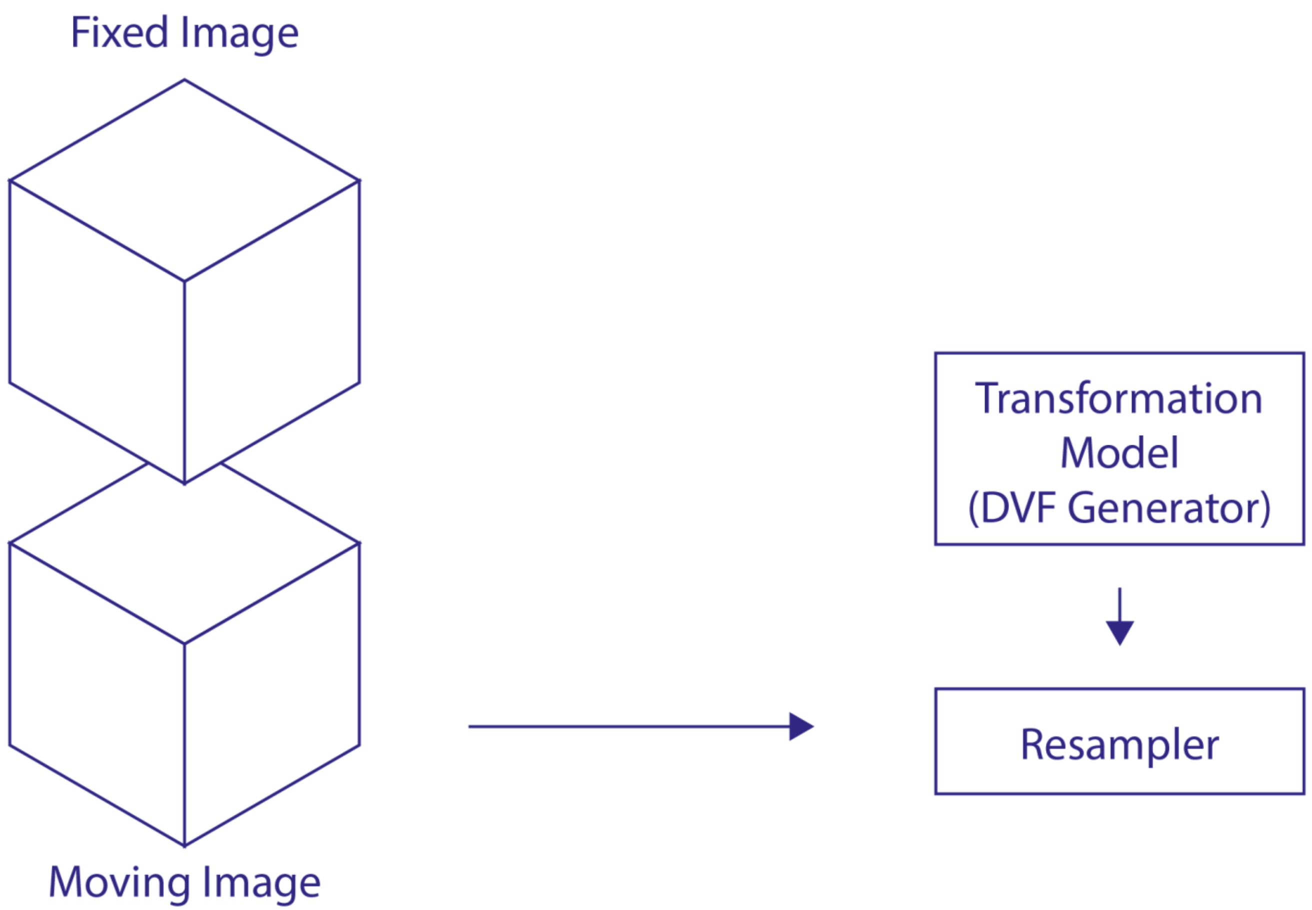


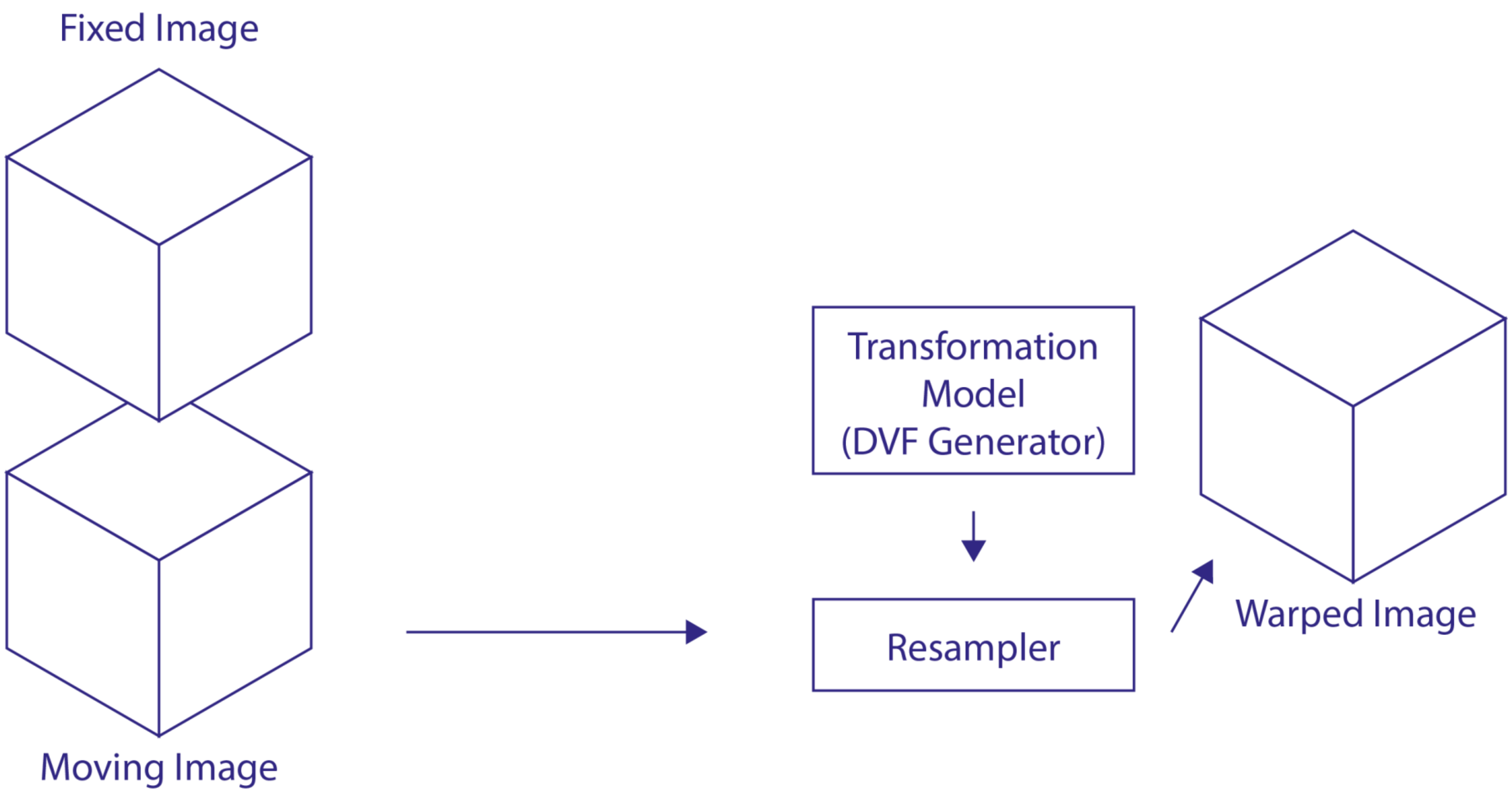
Fixed Image

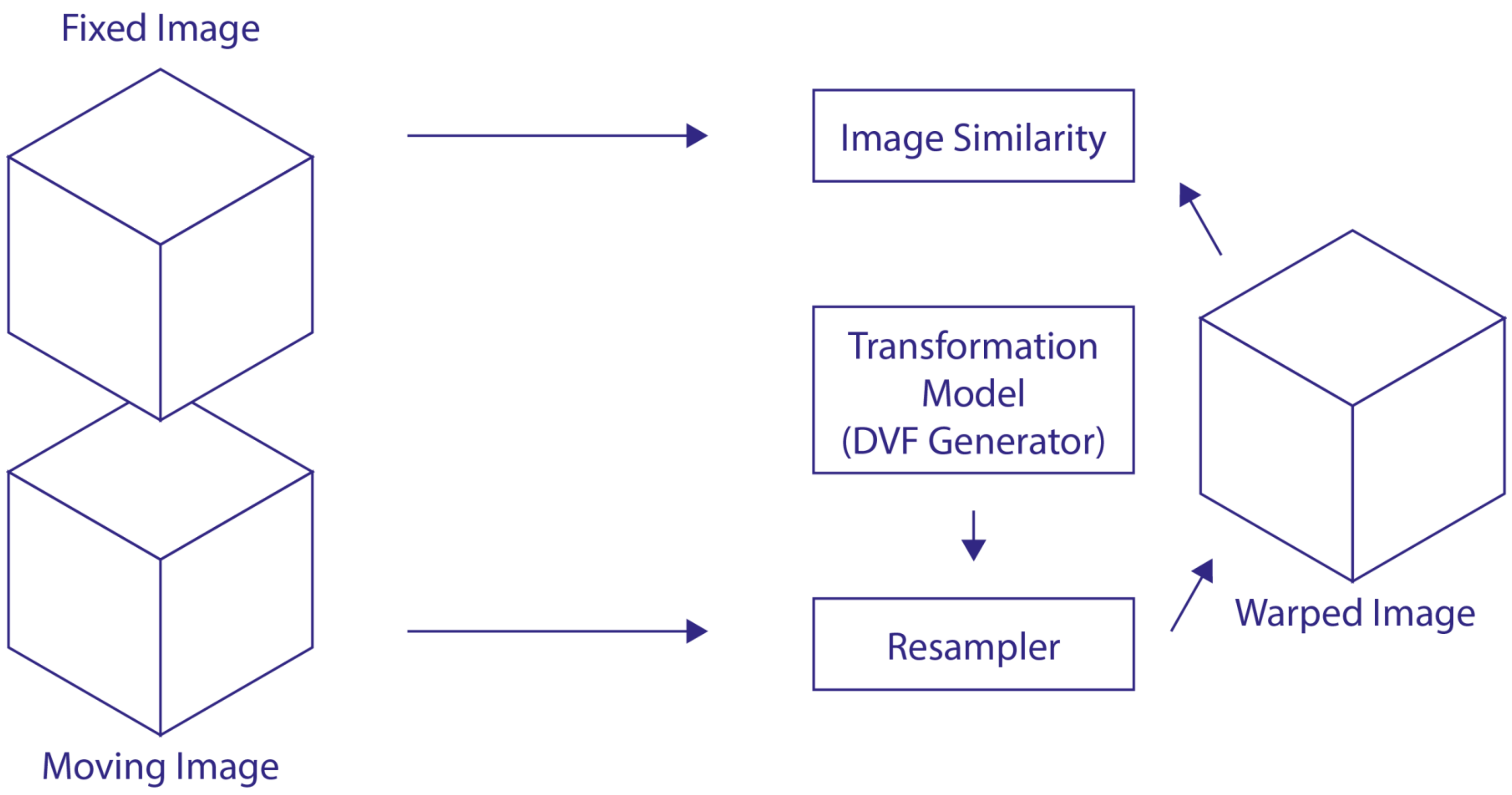


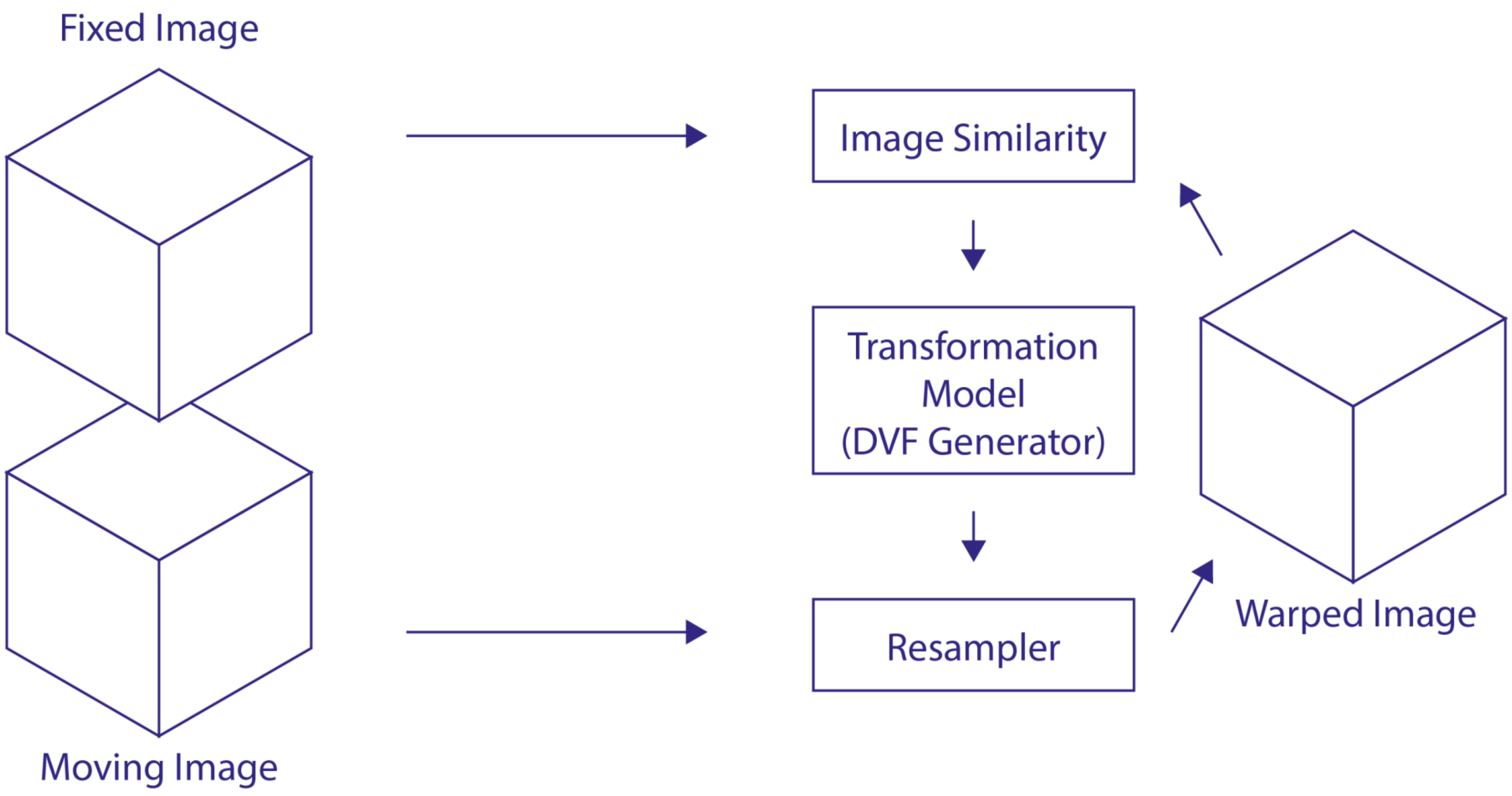
Moving Image

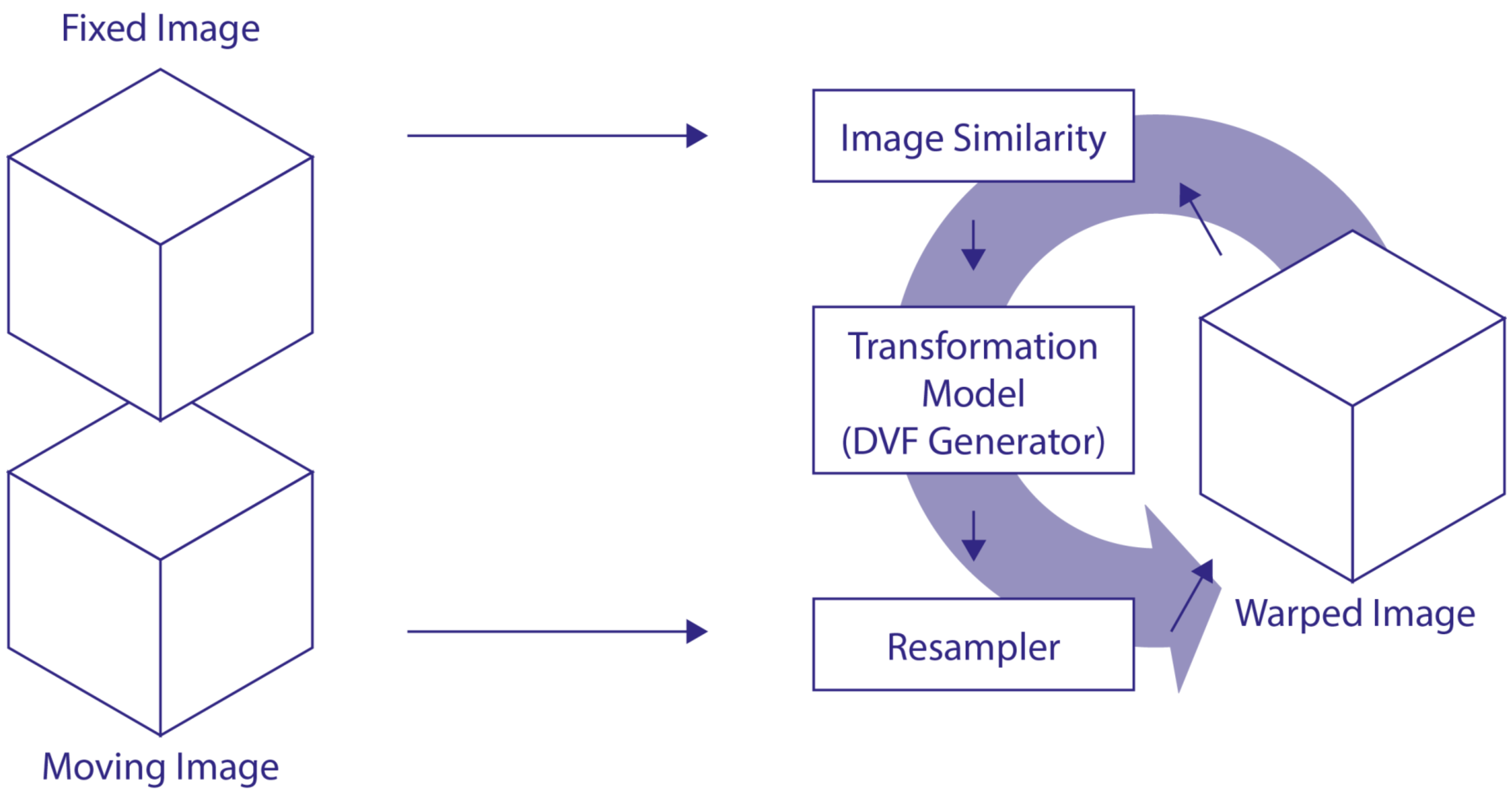
Transformation Model
(DVF Generator)

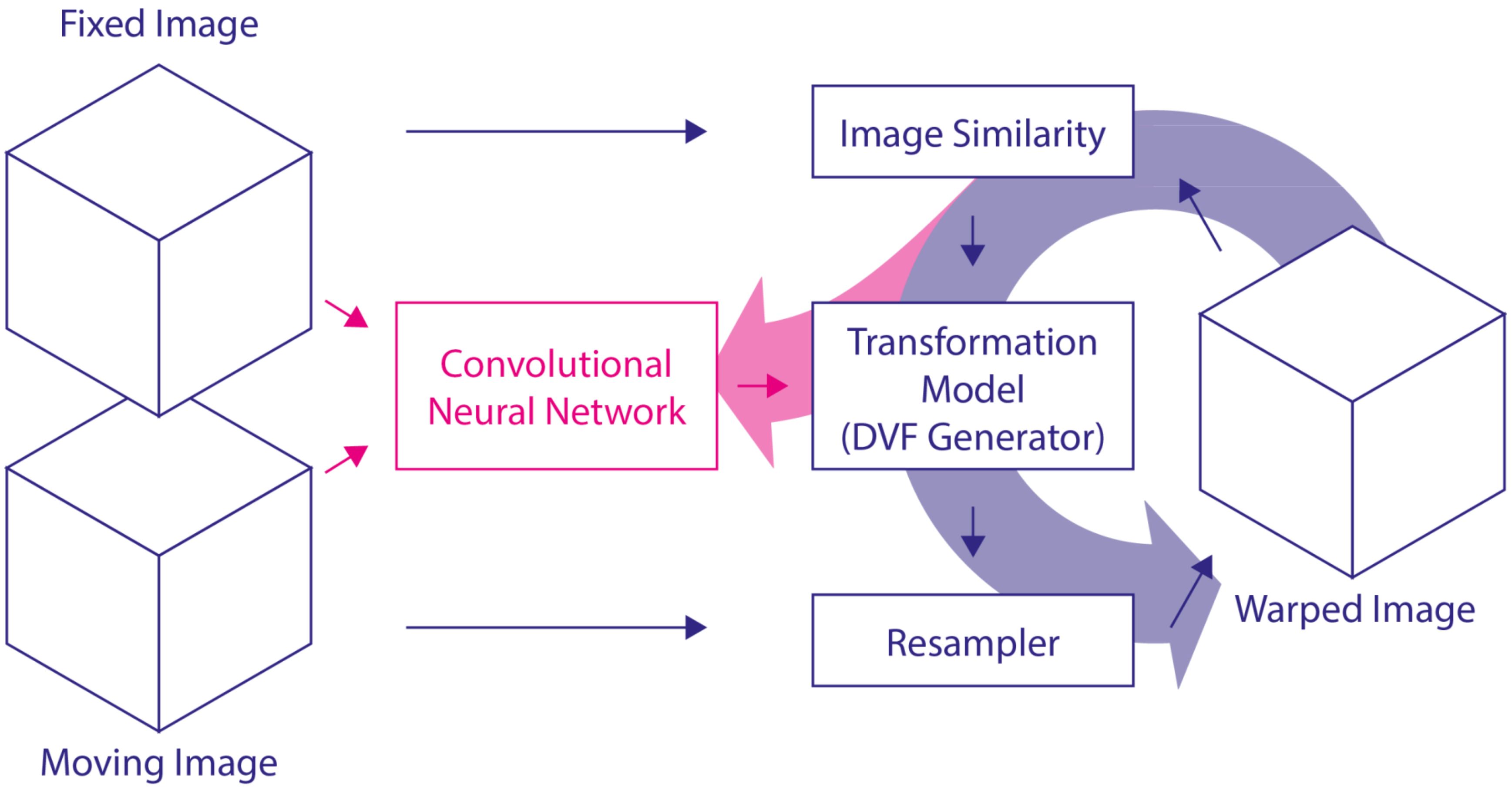

Fixed Image
Moving Image
Transformation Model (DVF Generator)
Resampler


Fixed Image
Moving Image
Transformation Model (DVF Generator)
Resampler
Warped Image


Fixed Image
Image Similarity
Transformation Model (DVF Generator)
Resampler
Warped Image
Moving Image


Fixed Image
Moving Image
Image Similarity
Transformation Model (DVF Generator)
Resampler
Warped Image


Fixed Image
Image Similarity
Transformation Model (DVF Generator)
Warped Image
Resampler
Moving Image



Fixed Image
Image Similarity
Convolutional Neural Network
Transformation Model (DVF Generator)
Warped Image
Resampler
Moving Image

MICCAI CONFERENCE 10.-14. SEPTEMBER 2017

DLMIA

3rd Workshop on Deep Learning in Medical Image Analysis

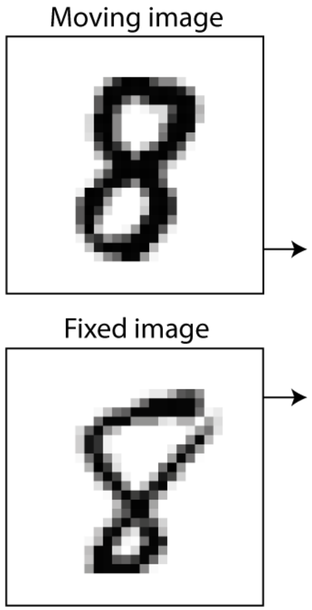
End-to-End Unsupervised Deformable Image Registration with a Convolutional Neural Network

Bob D. de Vos[1], Floris F. Berendsen[2], Max A. Viergever[1], Marius Staring[2], and Ivana Išgum[1]

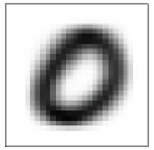
[1]Image Sciences Institute, University Medical Center Utrecht, the Netherlands
[2]Division of Image Processing, Leiden University Medical Center, the Netherlands

Image registration with unsupervised deep learning

Keywords: convolution neural network, deformable image registration, spatial transformer, cardiac MRI


Moving image
Fixed image

Average of 1,000 moving images from the test-set



Sunnybrook Cardiac Data: cardiac cine MRI

45 MR images

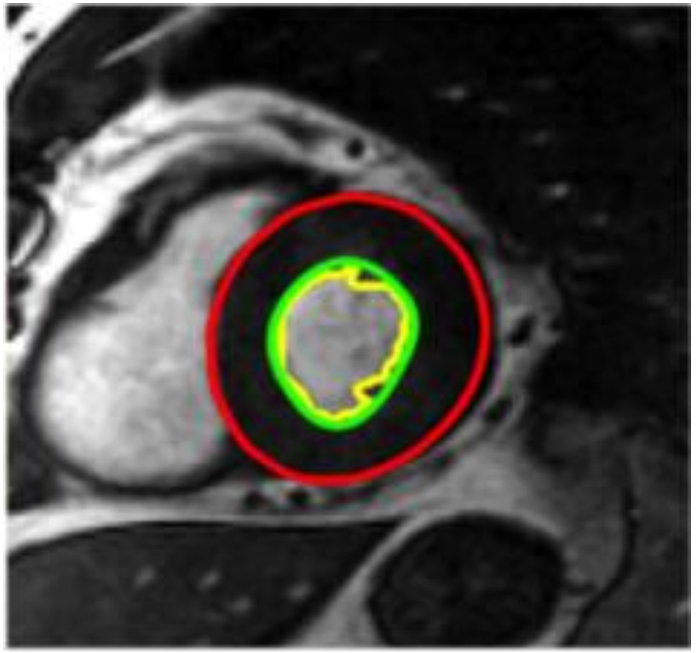
Training/validation/test→15/15/15

Image size 256x256x~10x20 voxels

Register 2D slices

from two time points

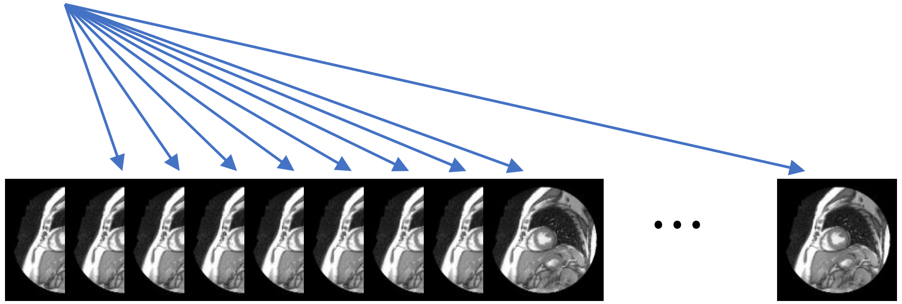
the same scan and slice location



Original

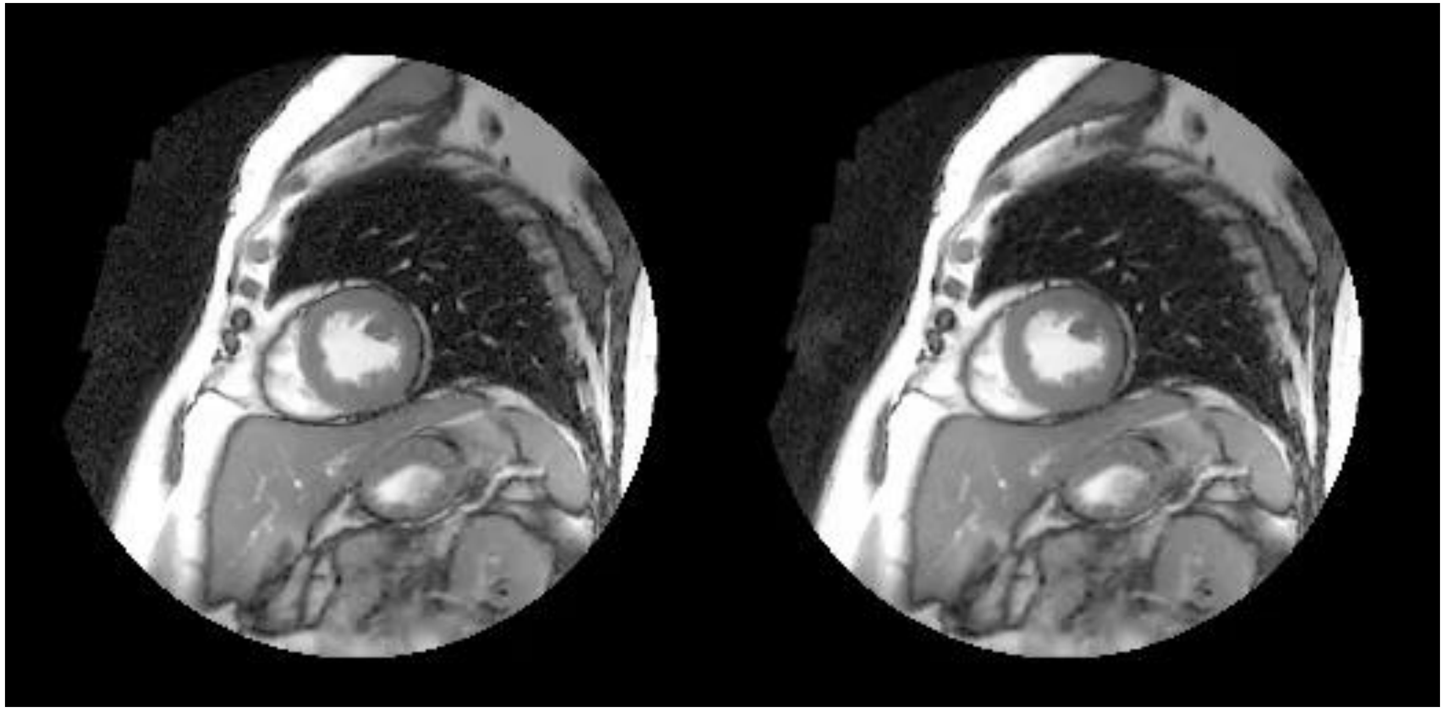
Original

Original

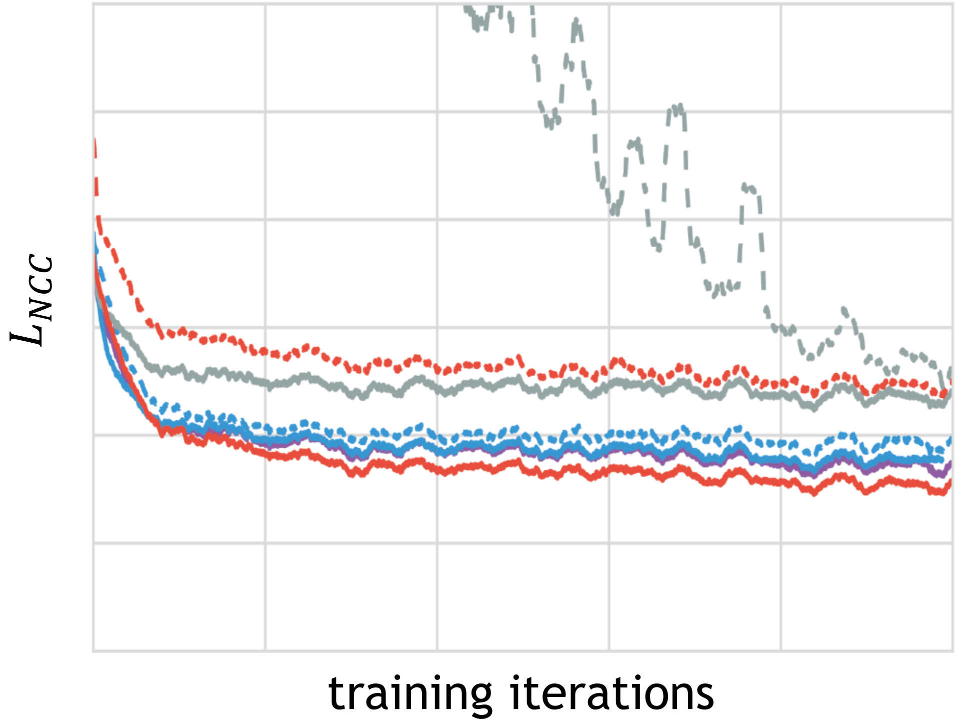


Original

Original Warped

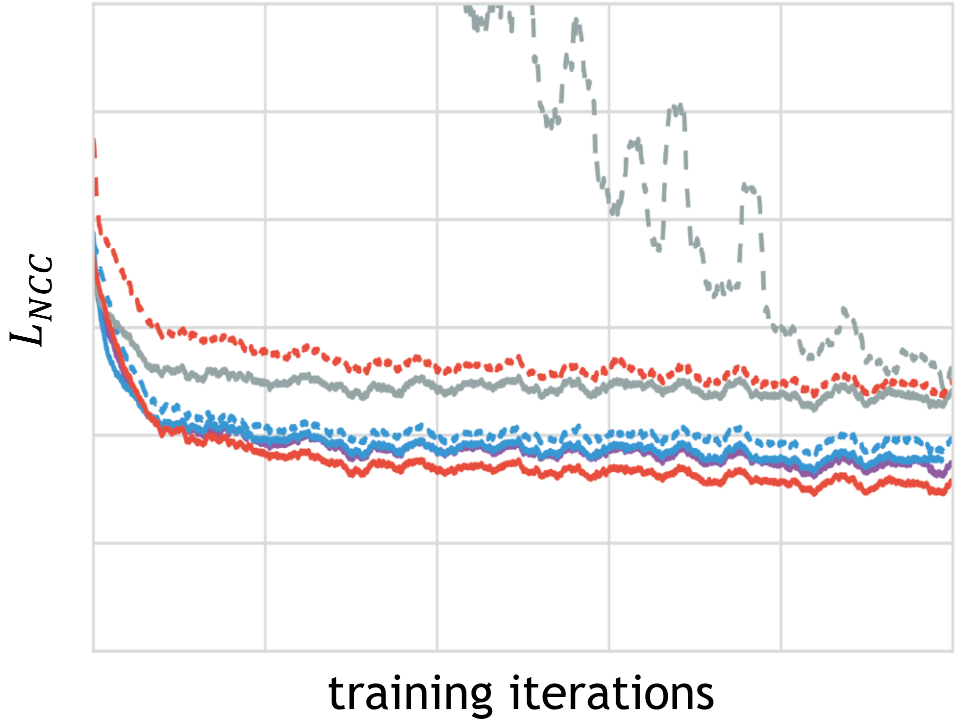


Original Warped Propagated labels



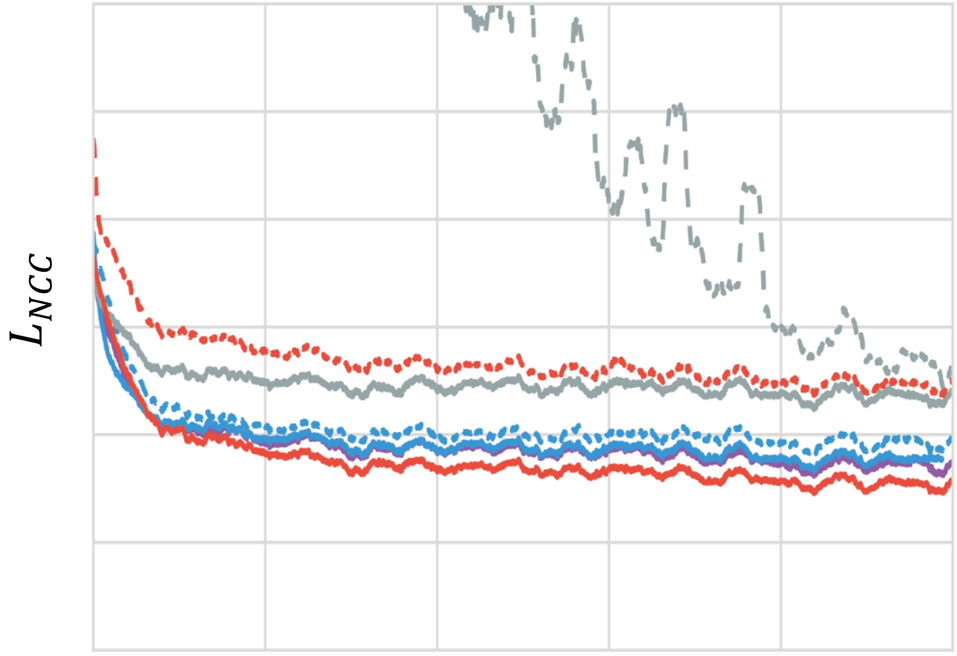
		Dice	95^{th}SD (mm)	MAD (mm)	Time (s)
No registration		0.62 ± 0.15	7.79 ± 2.92	2.89 ± 1.07	-
SimpleElastix	2×100	0.79 ± 0.08	5.09 ± 2.36	1.91 ± 0.94	0.51 ± 0.07
SimpleElastix	2×2000	$\mathbf{0.81 \pm 0.08}$	5.09 ± 7.25	$\mathbf{1.75 \pm 1.29}$	7.38 ± 0.94
Baseline		0.80 ± 0.08	$\mathbf{5.03 \pm 2.30}$	1.83 ± 0.89	0.049 ± 0.004
max pooling		0.78 ± 0.08	5.26 ± 2.16	1.95 ± 0.85	-
strided convolutions		0.78 ± 0.08	5.30 ± 2.28	1.97 ± 0.87	-
quadratic B-spline		0.72 ± 0.11	6.41 ± 2.61	2.40 ± 0.96	-
thin plate spline		0.78 ± 0.09	5.48 ± 2.36	2.01 ± 0.89	-
overlapping patches		0.79 ± 0.08	5.20 ± 2.30	1.92 ± 0.89	-
full image		0.76 ± 0.09	5.55 ± 2.24	2.10 ± 0.90	-

Evaluation: label propagation of left ventricle myocardium segmentation



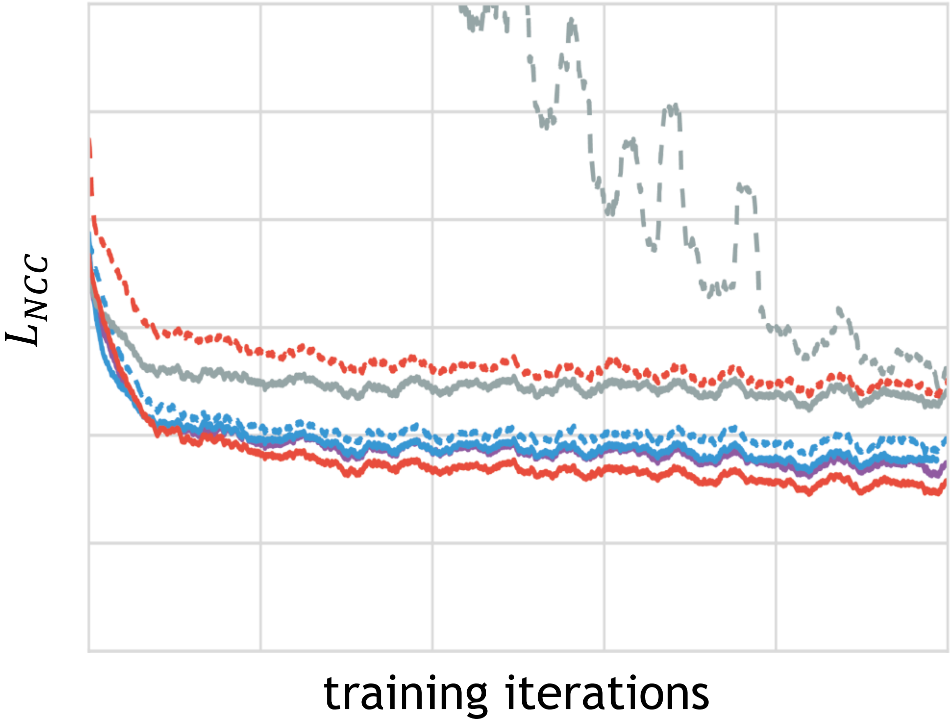
		Dice	95^{th}SD (mm)	MAD (mm)	Time (s)
No registration		0.62 ± 0.15	7.79 ± 2.92	2.89 ± 1.07	-
SimpleElastix	2×100	0.79 ± 0.08	5.09 ± 2.36	1.91 ± 0.94	0.51 ± 0.07
SimpleElastix	2×2000	$\mathbf{0.81 \pm 0.08}$	5.09 ± 7.25	$\mathbf{1.75 \pm 1.29}$	7.38 ± 0.94
Baseline		0.80 ± 0.08	$\mathbf{5.03 \pm 2.30}$	1.83 ± 0.89	0.049 ± 0.004
max pooling		0.78 ± 0.08	5.26 ± 2.16	1.95 ± 0.85	-
strided convolutions		0.78 ± 0.08	5.30 ± 2.28	1.97 ± 0.87	-
quadratic B-spline		0.72 ± 0.11	6.41 ± 2.61	2.40 ± 0.96	-
thin plate spline		0.78 ± 0.09	5.48 ± 2.36	2.01 ± 0.89	-
overlapping patches		0.79 ± 0.08	5.20 ± 2.30	1.92 ± 0.89	-
full image		0.76 ± 0.09	5.55 ± 2.24	2.10 ± 0.90	-

Evaluation: label propagation of left ventricle myocardium segmentation



		Dice	95^{th}SD (mm)	MAD (mm)	Time (s)
No registration		0.62 ± 0.15	7.79 ± 2.92	2.89 ± 1.07	-
SimpleElastix	2×100	0.79 ± 0.08	5.09 ± 2.36	1.91 ± 0.94	0.51 ± 0.07
SimpleElastix	2×2000	$\mathbf{0.81 \pm 0.08}$	5.09 ± 7.25	$\mathbf{1.75 \pm 1.29}$	7.38 ± 0.94
Baseline		0.80 ± 0.08	$\mathbf{5.03 \pm 2.30}$	1.83 ± 0.89	0.049 ± 0.004
max pooling		0.78 ± 0.08	5.26 ± 2.16	1.95 ± 0.85	-
strided convolutions		0.78 ± 0.08	5.30 ± 2.28	1.97 ± 0.87	-
quadratic B-spline		0.72 ± 0.11	6.41 ± 2.61	2.40 ± 0.96	-
thin plate spline		0.78 ± 0.09	5.48 ± 2.36	2.01 ± 0.89	-
overlapping patches		0.79 ± 0.08	5.20 ± 2.30	1.92 ± 0.89	-
full image		0.76 ± 0.09	5.55 ± 2.24	2.10 ± 0.90	-

Evaluation: label propagation of left ventricle myocardium segmentation



		Dice	95^{th}SD (mm)	MAD (mm)	Time (s)
No registration		0.62 ± 0.15	7.79 ± 2.92	2.89 ± 1.07	-
SimpleElastix	2×100	0.79 ± 0.08	5.09 ± 2.36	1.91 ± 0.94	0.51 ± 0.07
SimpleElastix	2×2000	**0.81 ± 0.08**	5.09 ± 7.25	**1.75 ± 1.29**	7.38 ± 0.94
Baseline		0.80 ± 0.08	**5.03 ± 2.30**	1.83 ± 0.89	0.049 ± 0.004
max pooling		0.78 ± 0.08	5.26 ± 2.16	1.95 ± 0.85	-
strided convolutions		0.78 ± 0.08	5.30 ± 2.28	1.97 ± 0.87	-
quadratic B-spline		0.72 ± 0.11	6.41 ± 2.61	2.40 ± 0.96	-
thin plate spline		0.78 ± 0.09	5.48 ± 2.36	2.01 ± 0.89	-
overlapping patches		0.79 ± 0.08	5.20 ± 2.30	1.92 ± 0.89	-
full image		0.76 ± 0.09	5.55 ± 2.24	2.10 ± 0.90	-

Evaluation: label propagation of left ventricle myocardium segmentation

Medical Image Analysis 52 (2019) 128–143

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

A deep learning framework for unsupervised affine and deformable image registration

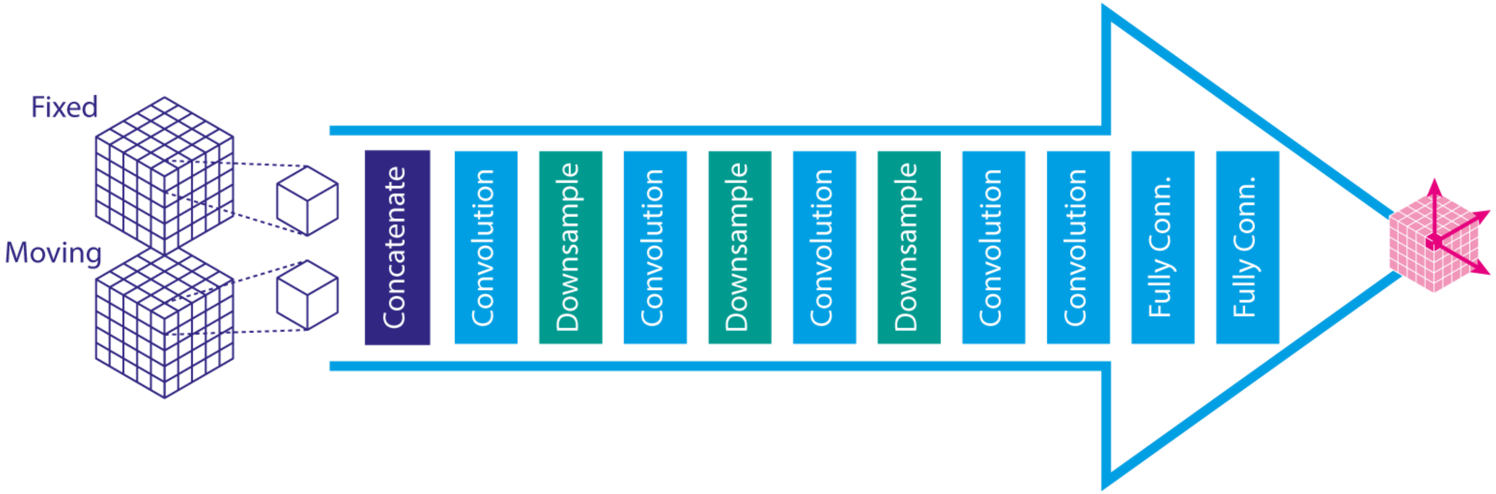
Bob D. de Vos[a,*], Floris F. Berendsen[b], Max A. Viergever[a], Hessam Sokooti[b], Marius Staring[b], Ivana Išgum[a]

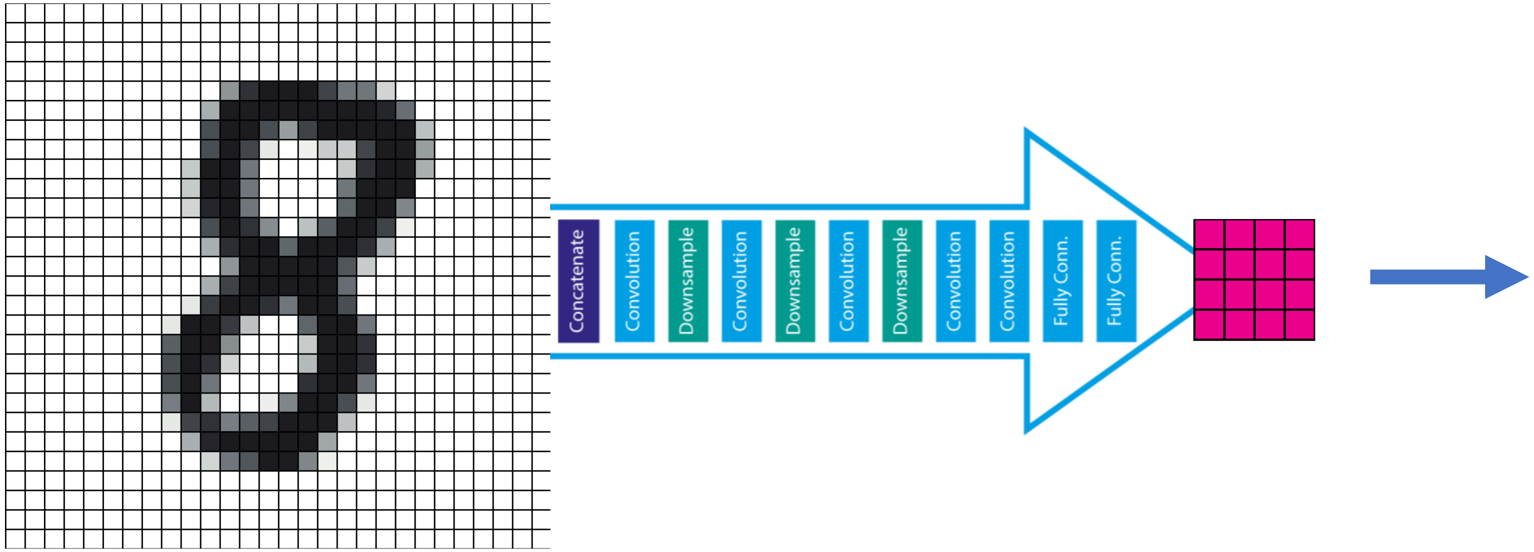
[a] Image Sciences Institute, University Medical Center Utrecht and Utrecht University, Utrecht, The Netherlands
[b] Division of Image Processing of the Leiden University Medical Center, Leiden, The Netherlands

Deep learning image registration framework

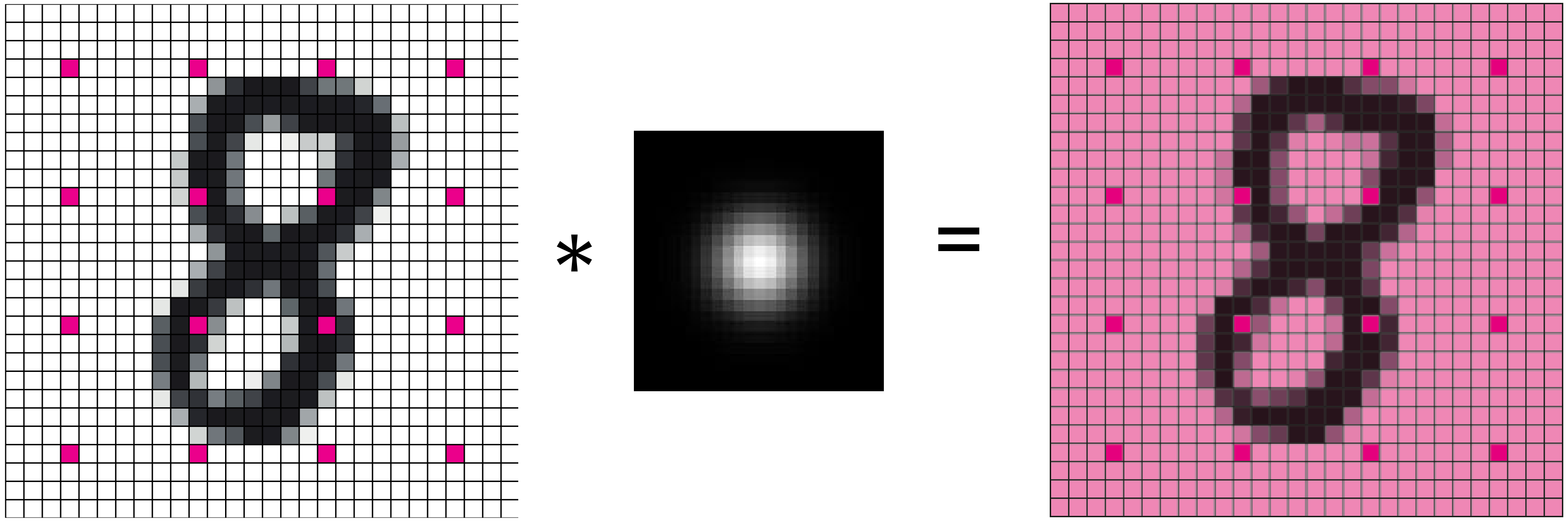
Deformable image registration

- 3D
- B-splines
- Patch-based
- Fully convolutional
- Bending energy penalty
- 3x3x3 convolutions
- 2x2x2 avg. pooling
- 32 kernels/nodes per layer

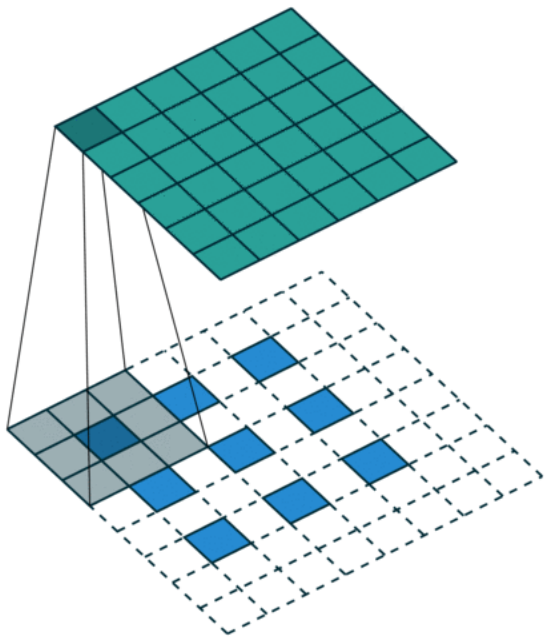



Concatenate
Convolution
Downsample
Convolution
Downsample
Convolution
Downsample
Convolution
Convolution
Fully Conn.
Fully Conn.

dense
DVF



- Transposed convolutions for b-spline registration
- Available in every DL package (except the b-spline kernel)

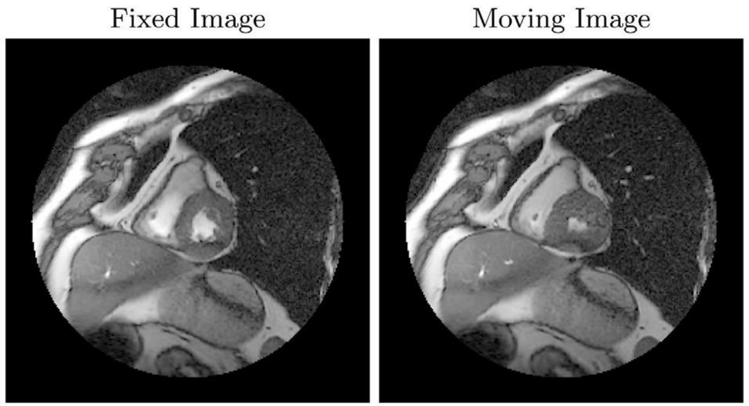


Regularization

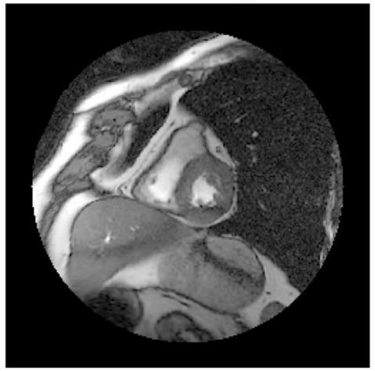
- Bending energy penalty

$$P = \frac{1}{V}\int_0^X \int_0^Y \int_0^Z \left[\left(\frac{\partial^2 \mathbf{T}}{\partial x^2}\right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial y^2}\right)^2 + \left(\frac{\partial^2 \mathbf{T}}{\partial z^2}\right)^2 + 2\left(\frac{\partial^2 \mathbf{T}}{\partial xy}\right)^2 + 2\left(\frac{\partial^2 \mathbf{T}}{\partial xz}\right)^2 + 2\left(\frac{\partial^2 \mathbf{T}}{\partial yz}\right)^2 \right] dx\, dy\, dz,$$

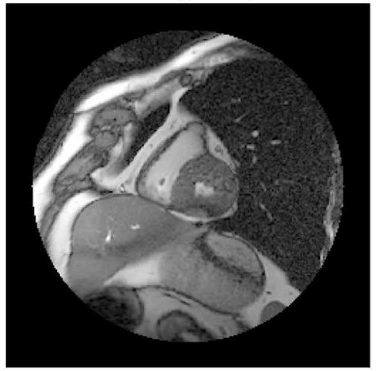
Rueckert et al., 1999, IEEE Trans. Med. Imag. 18 (8), 712–721.


Fixed Image
Moving Image

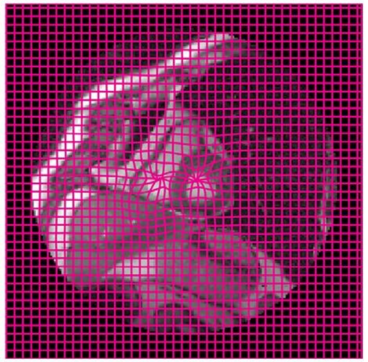
Fixed Image

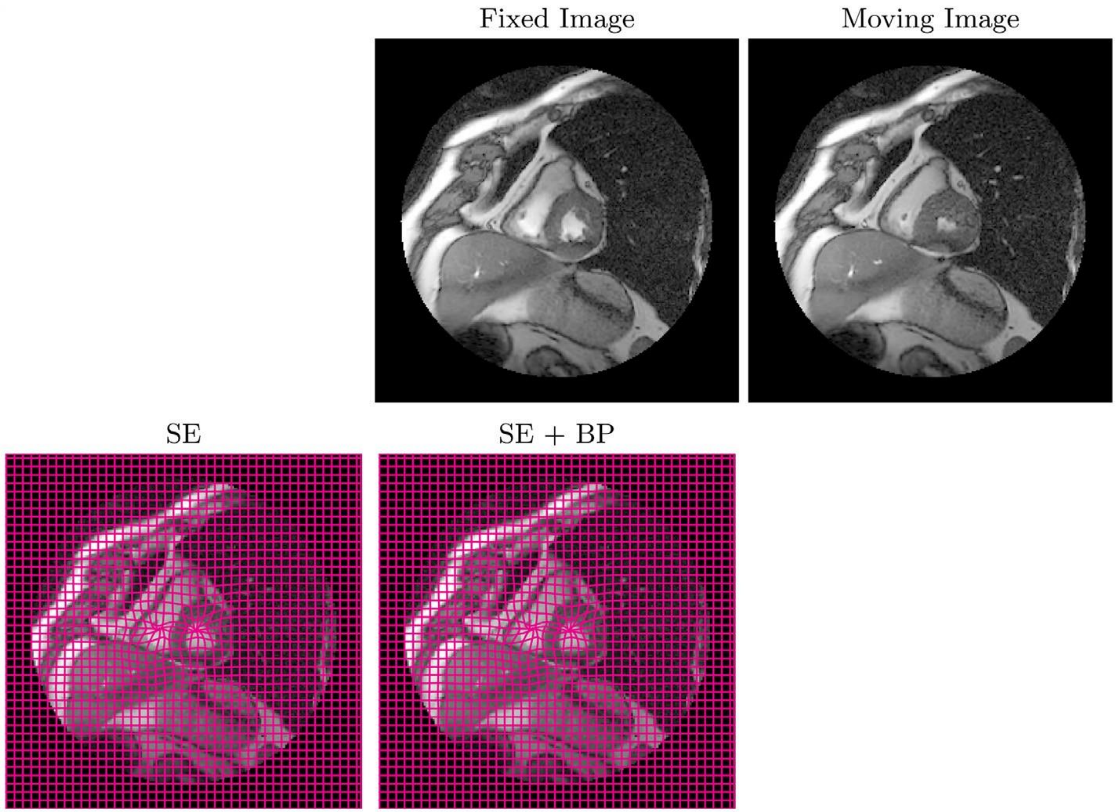


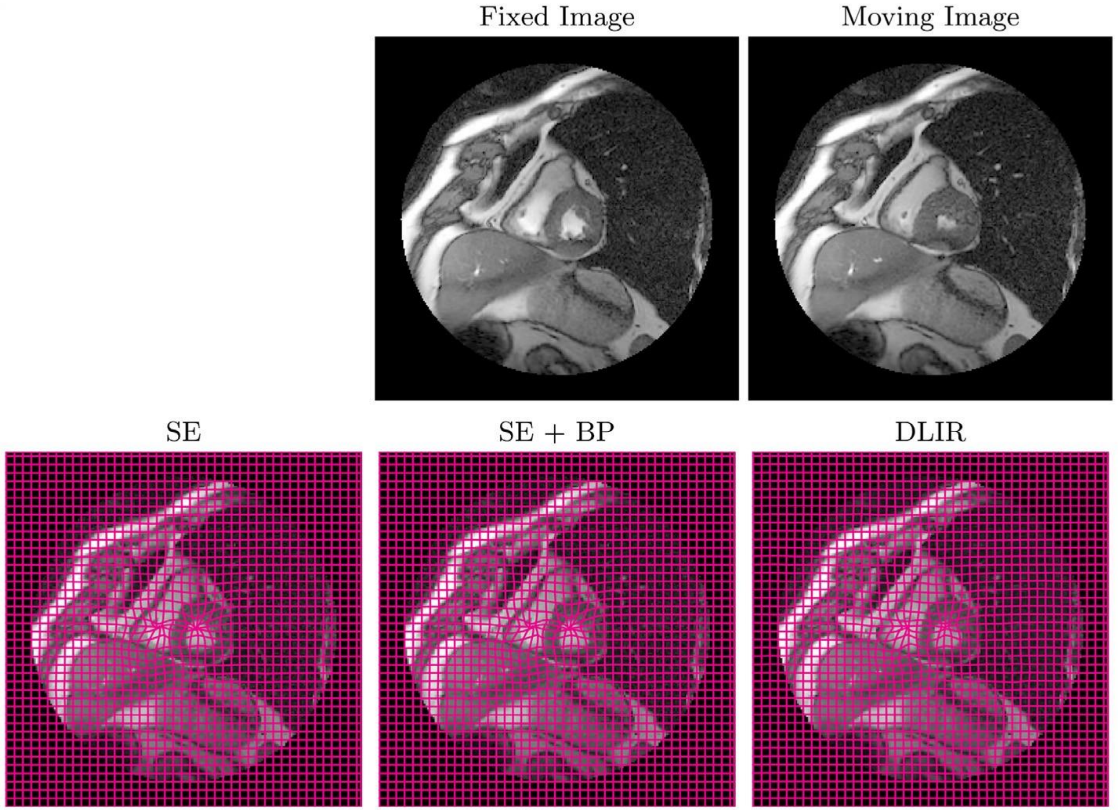
Moving Image

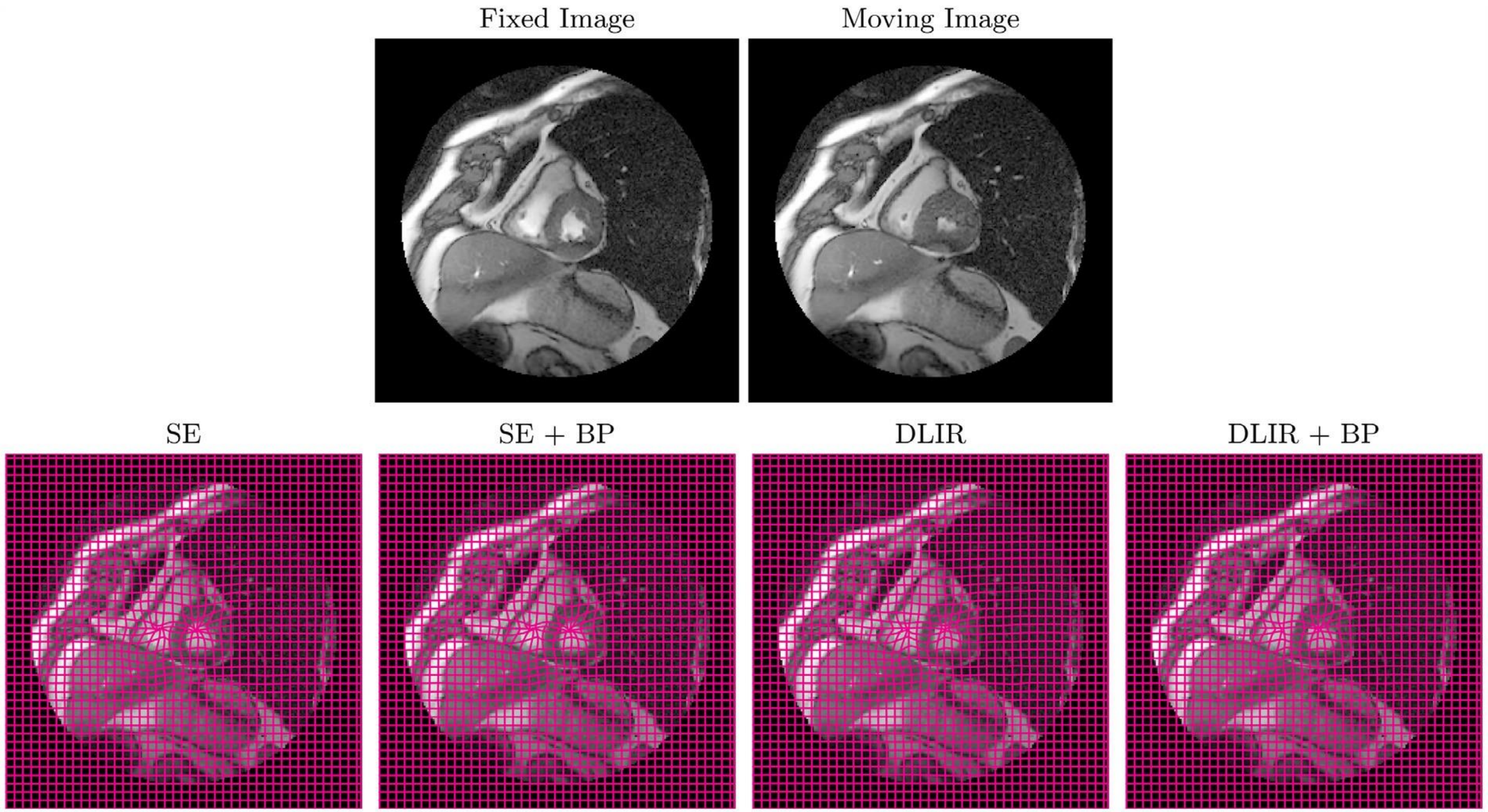


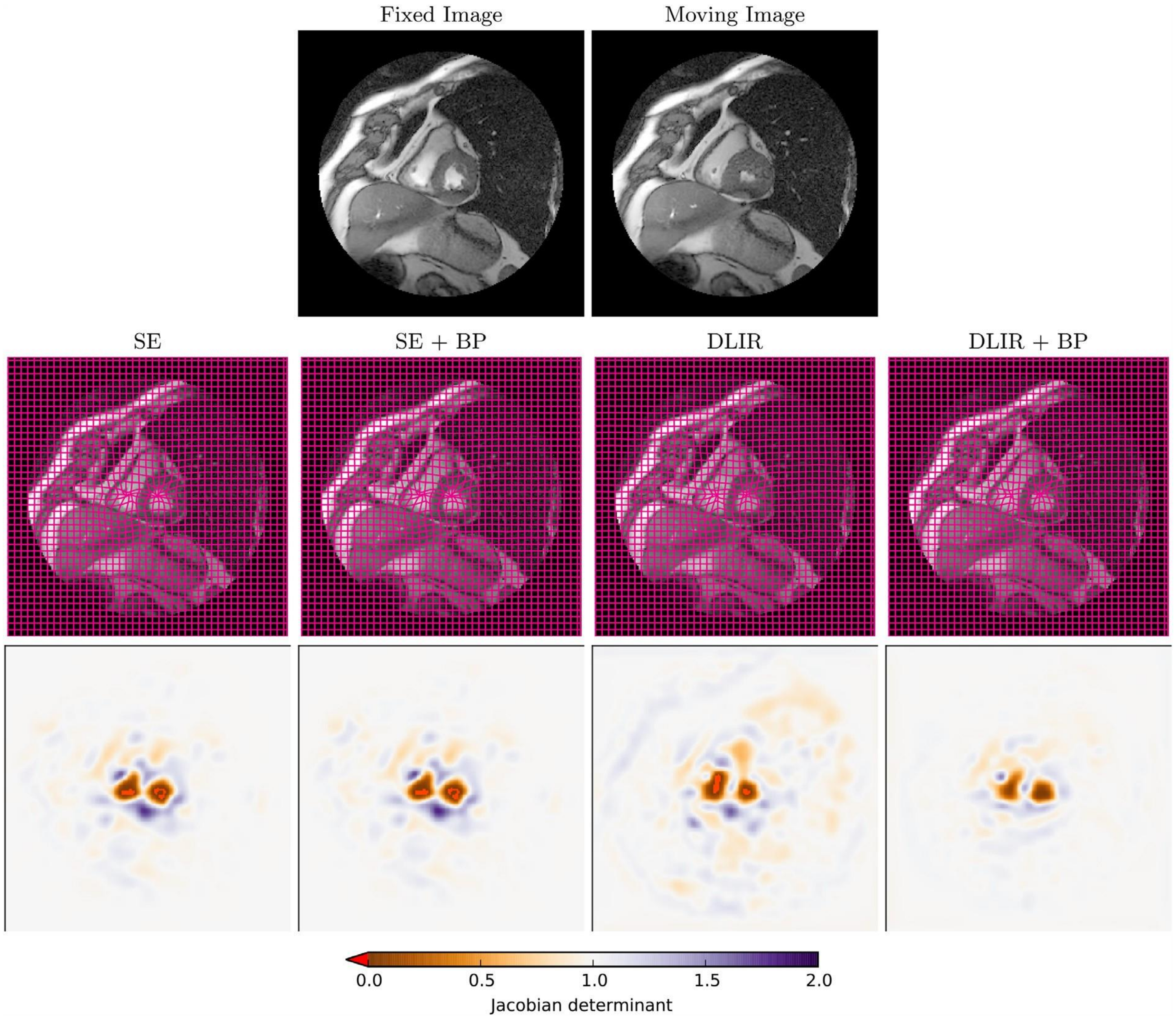
SE




Fixed Image
Moving Image
SE
SE + BP


Fixed Image
Moving Image
SE
SE + BP
DLIR


Fixed Image
Moving Image
SE
SE + BP
DLIR
DLIR + BP



Fixed Image
Moving Image
SE
SE + BP
DLIR
DLIR + BP
0.0
0.5
1.0
1.5
2.0
Jacobian determinant

		Dice	HD	ASD	Fraction folding	Std. dev. Jacobian	CPU time (s)	GPU time (s)
Before registration		0.70 ± 0.30	15.46 ± 4.50	4.66 ± 4.26	–	–	–	–
Single stage	SE	0.86 ± 0.18	9.76 ± 4.78	1.14 ± 1.40	0.08 ± 0.16	0.15 ± 0.08	13.49(3.27)	–
	SE + BP	0.86 ± 0.17	9.64 ± 4.15	1.13 ± 1.38	0.07 ± 0.15	0.15 ± 0.08	14.89(3.07)	–
	DLIR	0.87 ± 0.18	9.47 ± 5.26	0.98 ± 1.12	0.03 ± 0.06	0.14 ± 0.04	1.71 (0.45)	0.03 ± 0.01
	DLIR + BP	0.86 ± 0.18	9.10 ± 4.26	1.01 ± 1.42	0.00 ± 0.01	0.09 ± 0.03		

SE = conventional IR with SimpleElastix

DLIR = deep learning image registration

BP = bending energy penalty

		Dice	HD	ASD	Fraction folding	Std. dev. Jacobian	CPU time (s)	GPU time (s)
Before registration		0.70 ± 0.30	15.46 ± 4.50	4.66 ± 4.26	–	–	–	–
Single stage	SE	0.86 ± 0.18	9.76 ± 4.78	1.14 ± 1.40	0.08 ± 0.16	0.15 ± 0.08	13.49(3.27)	–
	SE + BP	0.86 ± 0.17	9.64 ± 4.15	1.13 ± 1.38	0.07 ± 0.15	0.15 ± 0.08	14.89(3.07)	–
	DLIR	0.87 ± 0.18	9.47 ± 5.26	0.98 ± 1.12	0.03 ± 0.06	0.14 ± 0.04	1.71 (0.45)	0.03 ± 0.01
	DLIR + BP	0.86 ± 0.18	9.10 ± 4.26	1.01 ± 1.42	0.00 ± 0.01	0.09 ± 0.03		

SE = conventional IR with SimpleElastix

DLIR = deep learning image registration

BP = bending energy penalty

		Dice	HD	ASD	Fraction folding	Std. dev. Jacobian	CPU time (s)	GPU time (s)
Before registration		0.70 ± 0.30	15.46 ± 4.50	4.66 ± 4.26	–	–	–	–
Single stage	SE	0.86 ± 0.18	9.76 ± 4.78	1.14 ± 1.40	0.08 ± 0.16	0.15 ± 0.08	13.49(3.27)	–
	SE + BP	0.86 ± 0.17	9.64 ± 4.15	1.13 ± 1.38	0.07 ± 0.15	0.15 ± 0.08	14.89(3.07)	–
	DLIR	0.87 ± 0.18	9.47 ± 5.26	0.98 ± 1.12	0.03 ± 0.06	0.14 ± 0.04	1.71 (0.45)	0.03 ± 0.01
	DLIR + BP	0.86 ± 0.18	9.10 ± 4.26	1.01 ± 1.42	0.00 ± 0.01	0.09 ± 0.03		

SE = conventional IR with SimpleElastix

DLIR = deep learning image registration

BP = bending energy penalty

		Dice	HD	ASD	Fraction folding	Std. dev. Jacobian	CPU time (s)	GPU time (s)
Before registration		0.70 ± 0.30	15.46 ± 4.50	4.66 ± 4.26	–	–	–	–
Single stage	SE	0.86 ± 0.18	9.76 ± 4.78	1.14 ± 1.40	0.08 ± 0.16	0.15 ± 0.08	13.49(3.27)	–
	SE + BP	0.86 ± 0.17	9.64 ± 4.15	1.13 ± 1.38	0.07 ± 0.15	0.15 ± 0.08	14.89(3.07)	–
	DLIR	0.87 ± 0.18	9.47 ± 5.26	0.98 ± 1.12	0.03 ± 0.06	0.14 ± 0.04	1.71 (0.45)	0.03 ± 0.01
	DLIR + BP	0.86 ± 0.18	9.10 ± 4.26	1.01 ± 1.42	0.00 ± 0.01	0.09 ± 0.03		

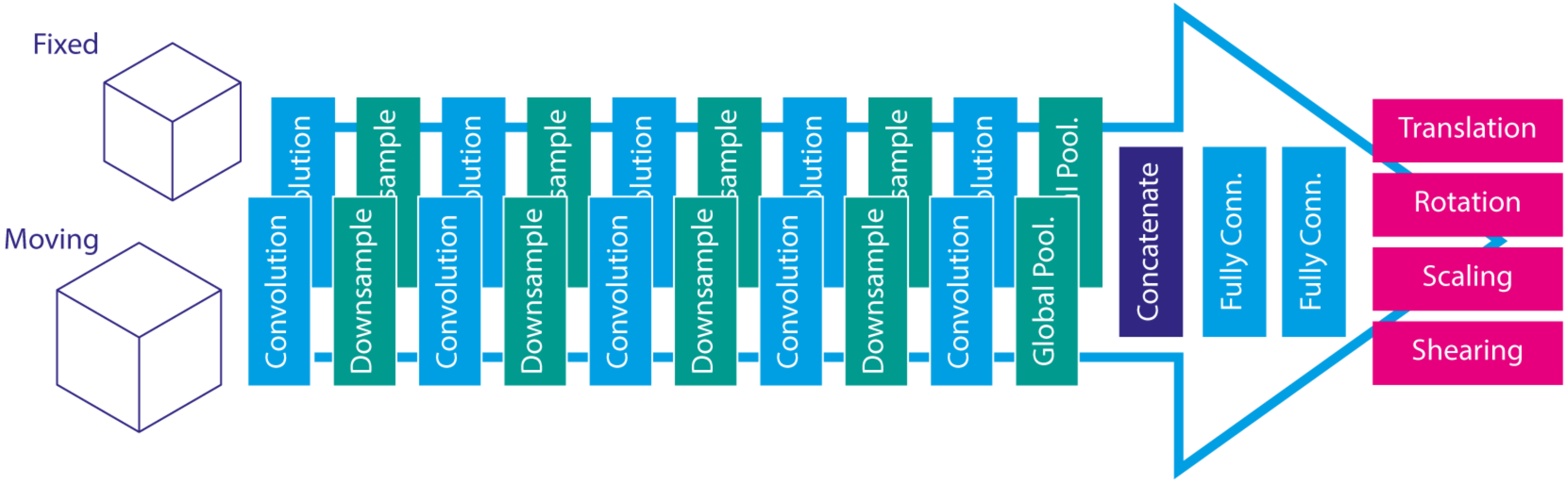
SE = conventional IR with SimpleElastix

DLIR = deep learning image registration

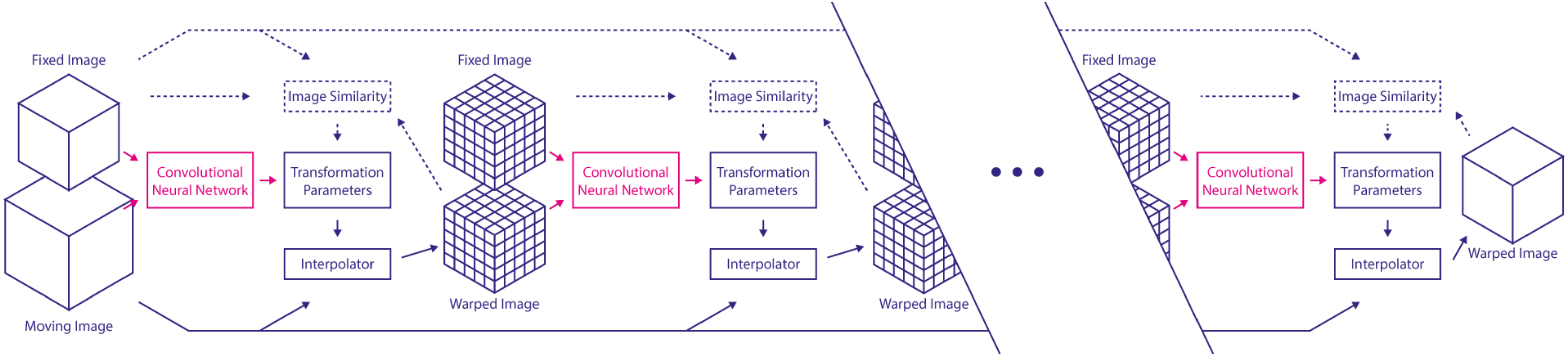
BP = bending energy penalty

Global image registration

- Affine registration
- Different size input
- 3x3x3 convolutions
- 2x2x2 avg. Pooling
- 32 kernels per layer



Multi-stage image registration



Global image registration (e.g. affine)

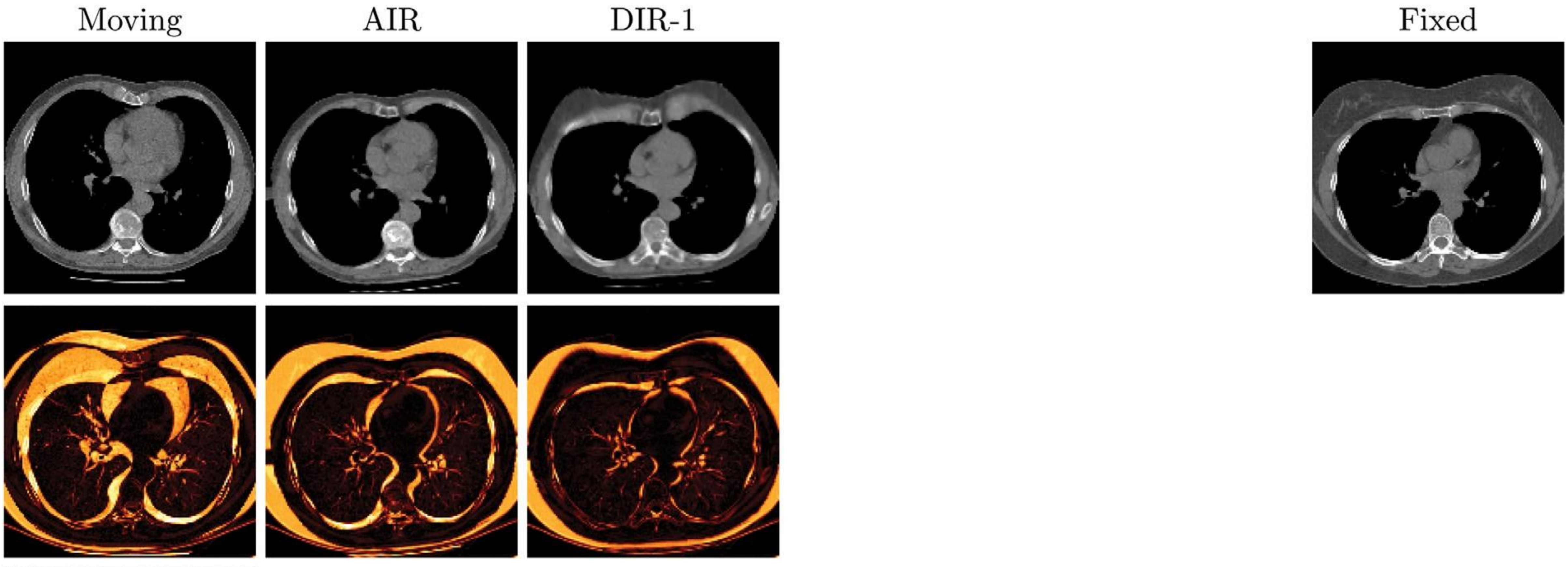
Coarse-to-fine image registration

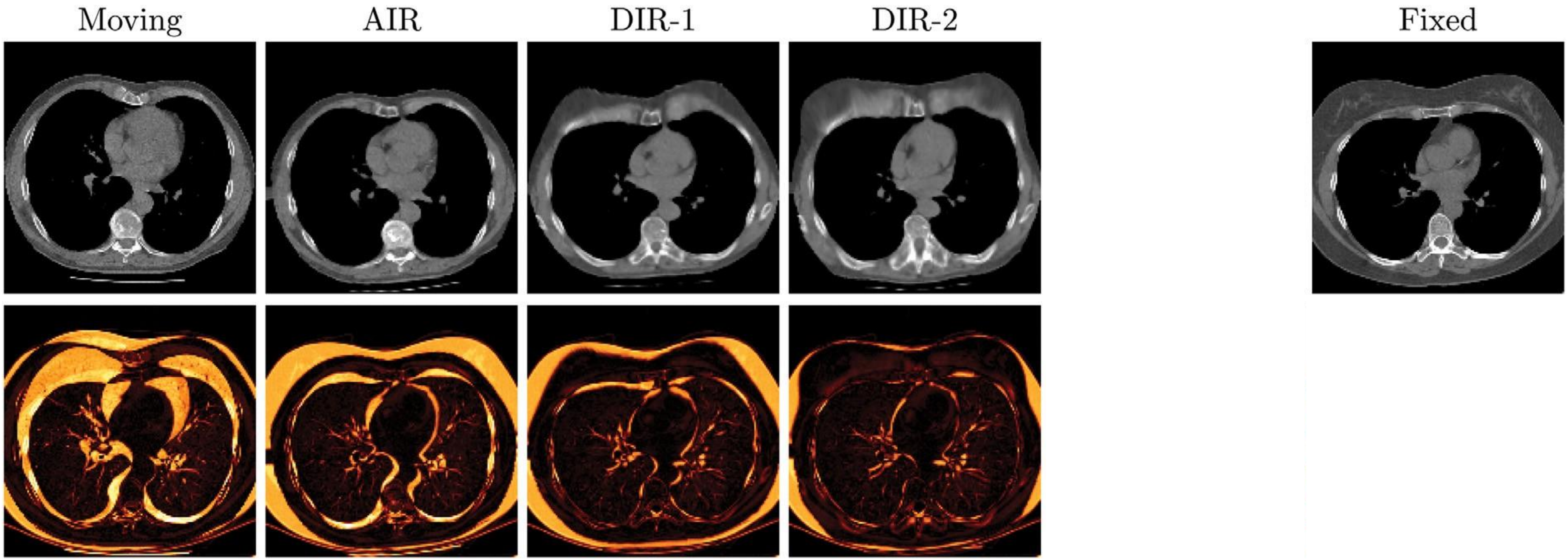
Schnabel et al., MICCAI 2001, pp. 573-581

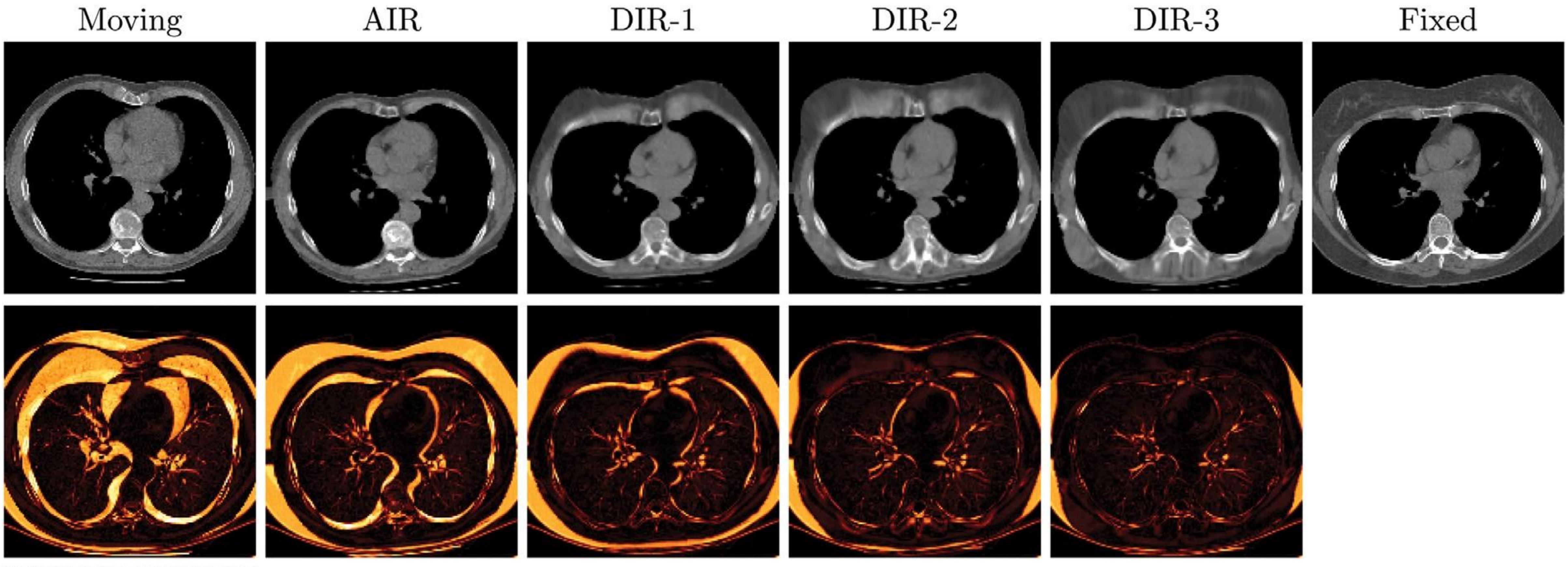
Moving

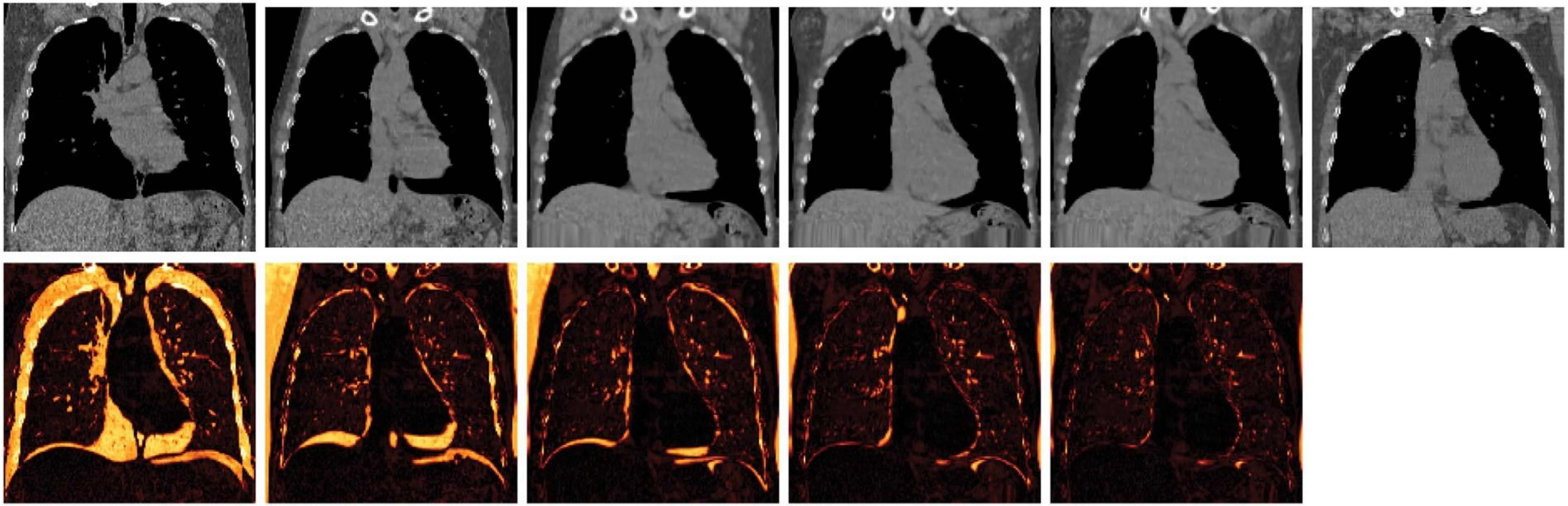
Fixed

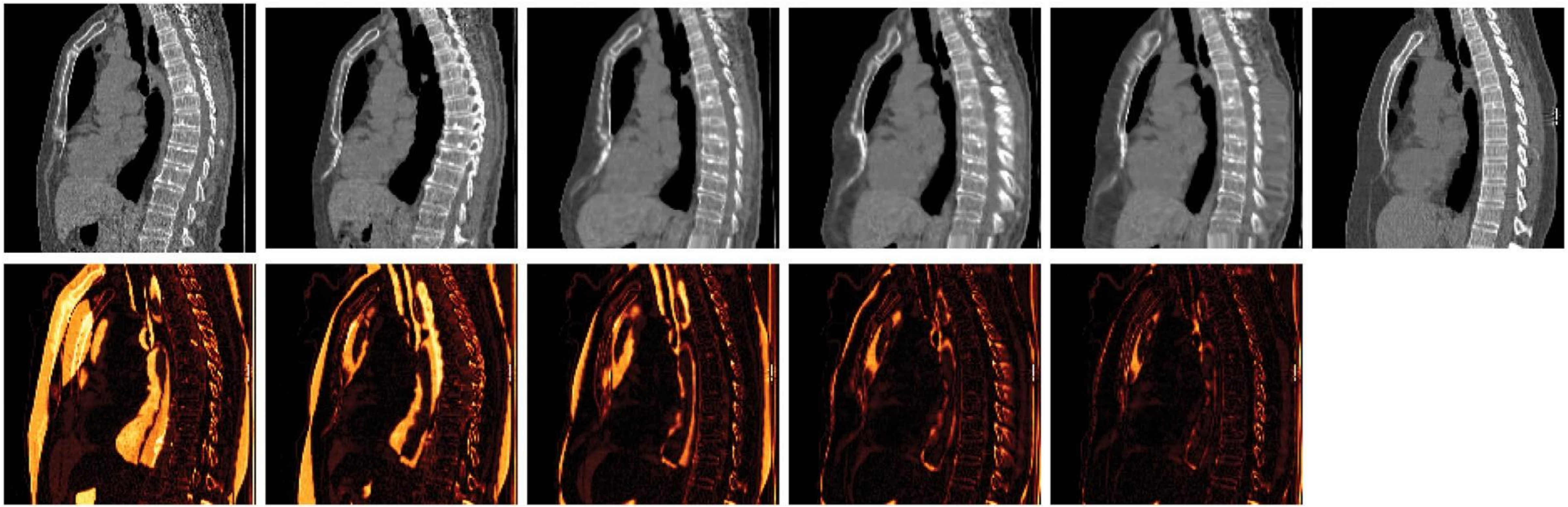

Moving
AIR
Fixed

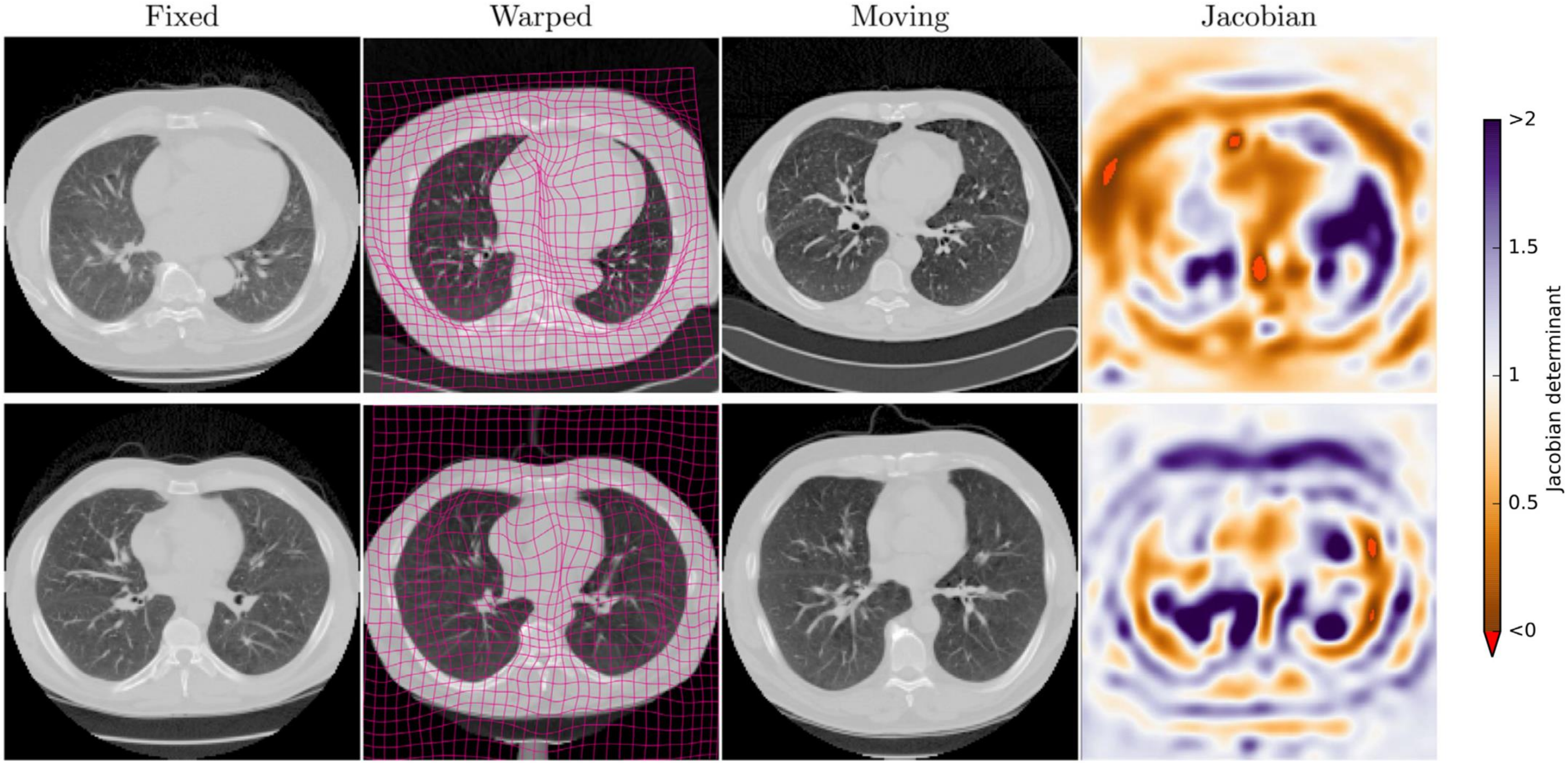

Moving
AIR
DIR-1
Fixed


Moving
AIR
DIR-1
DIR-2
Fixed


Moving
AIR
DIR-1
DIR-2
DIR-3
Fixed






Fixed
Warped
Moving
Jacobian
>2
1.5
1
0.5
<0
Jacobian determinant

		Dice	HD	ASD	Fraction folding (%)	Std. dev. Jacobian	CPU time (s)	GPU time (s)
Before registration		0.31 ± 0.21	32.62 ± 12.21	9.21 ± 4.53	–	–	–	–
SE	AIR	0.60 ± 0.19	25.81 ± 15.34	4.89 ± 2.36	–	–	3.73(0.26)	–
	DIR-1	0.69 ± 0.11	20.30 ± 13.26	3.39 ± 1.11	0.00 ± 0.00	0.19 ± 0.11	11.67(1.07)	–
	DIR-2	0.75 ± 0.08	21.26 ± 11.31	2.67 ± 0.87	0.00 ± 0.08	0.27 ± 0.13	14.83(3.37)	–
	DIR-3	0.77 ± 0.08	20.83 ± 11.81	2.45 ± 0.89	0.04 ± 0.19	0.30 ± 0.15	20.36(8.41)	–
DLIR	AIR	0.58 ± 0.16	26.79 ± 13.05	5.24 ± 2.19	–	–	1.02(0.29)	0.17(0.05)
	DIR-1	0.64 ± 0.11	21.68 ± 13.09	3.86 ± 1.74	0.00 ± 0.00	0.16 ± 0.09	3.85(0.99)	0.18(0.05)
	DIR-2	0.70 ± 0.10	19.95 ± 13.30	3.21 ± 1.15	0.00 ± 0.00	0.19 ± 0.10	8.18(2.03)	0.30(0.07)
	DIR-3	0.75 ± 0.08	19.34 ± 13.41	2.46 ± 0.80	0.75 ± 1.08	0.45 ± 0.21	15.41(4.38)	0.43(0.10)

		Dice	HD	ASD	Fraction folding (%)	Std. dev. Jacobian	CPU time (s)	GPU time (s)
Before registration		0.31 ± 0.21	32.62 ± 12.21	9.21 ± 4.53	–	–	–	–
SE	AIR	0.60 ± 0.19	25.81 ± 15.34	4.89 ± 2.36	–	–	3.73(0.26)	–
	DIR-1	0.69 ± 0.11	20.30 ± 13.26	3.39 ± 1.11	0.00 ± 0.00	0.19 ± 0.11	11.67(1.07)	–
	DIR-2	0.75 ± 0.08	21.26 ± 11.31	2.67 ± 0.87	0.00 ± 0.08	0.27 ± 0.13	14.83(3.37)	–
	DIR-3	0.77 ± 0.08	20.83 ± 11.81	2.45 ± 0.89	0.04 ± 0.19	0.30 ± 0.15	20.36(8.41)	–
DLIR	AIR	0.58 ± 0.16	26.79 ± 13.05	5.24 ± 2.19	–	–	1.02(0.29)	0.17(0.05)
	DIR-1	0.64 ± 0.11	21.68 ± 13.09	3.86 ± 1.74	0.00 ± 0.00	0.16 ± 0.09	3.85(0.99)	0.18(0.05)
	DIR-2	0.70 ± 0.10	19.95 ± 13.30	3.21 ± 1.15	0.00 ± 0.00	0.19 ± 0.10	8.18(2.03)	0.30(0.07)
	DIR-3	0.75 ± 0.08	19.34 ± 13.41	2.46 ± 0.80	0.75 ± 1.08	0.45 ± 0.21	15.41(4.38)	0.43(0.10)

		Dice	HD	ASD	Fraction folding (%)	Std. dev. Jacobian	CPU time (s)	GPU time (s)
Before registration		0.31 ± 0.21	32.62 ± 12.21	9.21 ± 4.53	–	–	–	–
SE	AIR	0.60 ± 0.19	25.81 ± 15.34	4.89 ± 2.36	–	–	3.73(0.26)	–
	DIR-1	0.69 ± 0.11	20.30 ± 13.26	3.39 ± 1.11	0.00 ± 0.00	0.19 ± 0.11	11.67(1.07)	–
	DIR-2	0.75 ± 0.08	21.26 ± 11.31	2.67 ± 0.87	0.00 ± 0.08	0.27 ± 0.13	14.83(3.37)	–
	DIR-3	0.77 ± 0.08	20.83 ± 11.81	2.45 ± 0.89	0.04 ± 0.19	0.30 ± 0.15	20.36(8.41)	–
DLIR	AIR	0.58 ± 0.16	26.79 ± 13.05	5.24 ± 2.19	–	–	1.02(0.29)	0.17(0.05)
	DIR-1	0.64 ± 0.11	21.68 ± 13.09	3.86 ± 1.74	0.00 ± 0.00	0.16 ± 0.09	3.85(0.99)	0.18(0.05)
	DIR-2	0.70 ± 0.10	19.95 ± 13.30	3.21 ± 1.15	0.00 ± 0.00	0.19 ± 0.10	8.18(2.03)	0.30(0.07)
	DIR-3	0.75 ± 0.08	19.34 ± 13.41	2.46 ± 0.80	0.75 ± 1.08	0.45 ± 0.21	15.41(4.38)	0.43(0.10)

Evaluation on DIRLab data

- Inspiration-expiration chest CT
- Evaluation: TRE of manually annotated landmarks
- Coarse-to-fine image registration:
  - Stage 1: 5.12 (4.64) mm
  - Stage 2: 3.40 (4.17) mm
  - Stage 3: 2.64 (4.32) mm

IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 38, NO. 9, SEPTEMBER 2019 2127

Direct Automatic Coronary Calcium Scoring in Cardiac and Chest CT

Bob D. de Vos, Jelmer M. Wolterink, Tim Leiner, Pim A. de Jong,
Nikolas Lessmann, and Ivana Išgum

Abstract—**Cardiovascular disease (CVD) is the global leading cause of death. A strong risk factor for CVD events is the amount of coronary artery calcium (CAC). To meet the demands of the increasing interest in quantification of CAC, i.e., coronary calcium scoring, especially as an unrequested finding for screening and research, automatic methods have been proposed. The current automatic cal-**

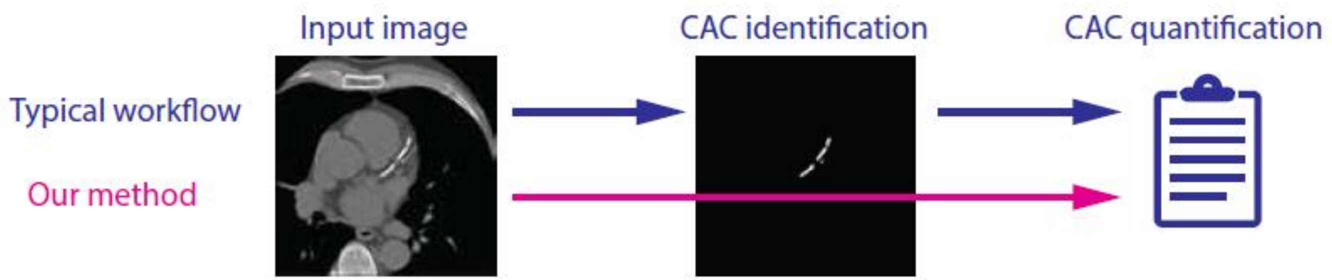
Group-wise image registration

weighted Cohen's kappa of 0.95 in cardiac CT and 0.93 in chest CT. Performance is similar to that of the state-of-the-art methods, but the proposed method is hundreds of times faster. By providing visual feedback, insight is given in the decision process, making it readily implementable in clinical

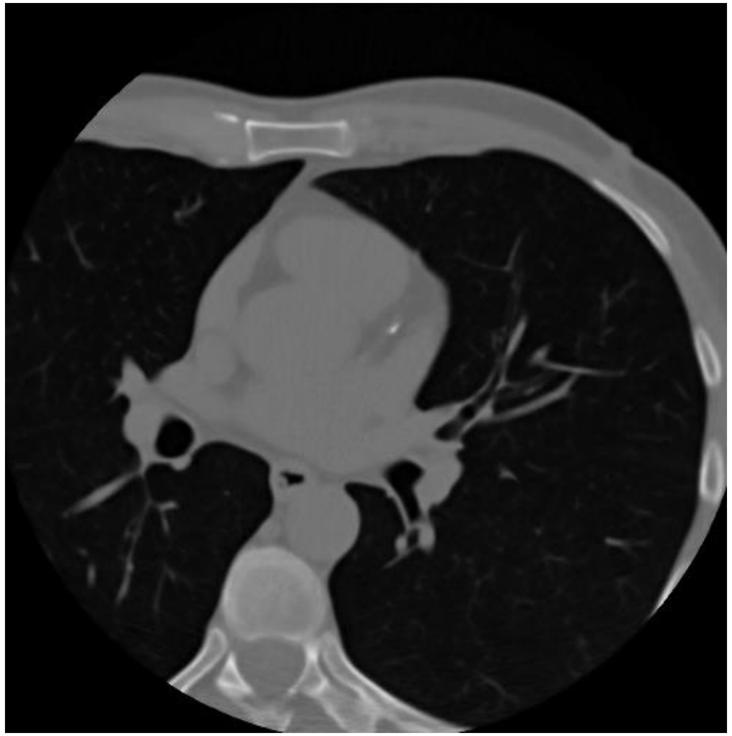
need for early detection and treatment of individuals with CVD or those who are at high cardiovascular risk due to the presence of one or more risk factors [2]. A strong and independent risk factor for CVD events, e.g. myocardial infarction, is the quantity of coronary artery calcium (CAC) [3]–[5]. Quantification of CAC, i.e. calcium scoring, is typically per-

manually identify CAC in CT image slices. This is a tedious process of finding and selecting high density voxels in the coronary arteries—commonly defined as two or more connected voxels above 130 Hounsfield Units (HU). In scans not dedicated to calcium scoring, this can be particularly

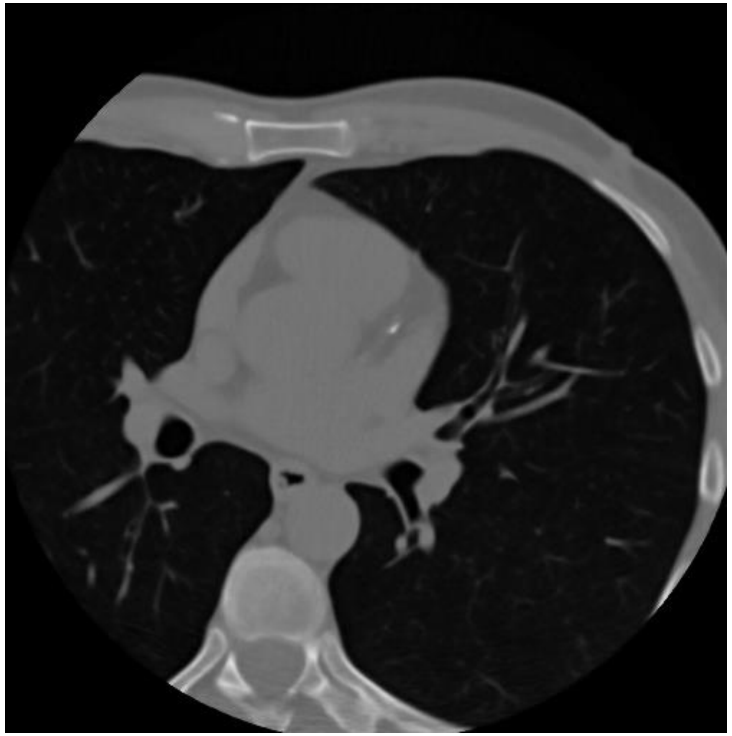
Direct calcium scoring



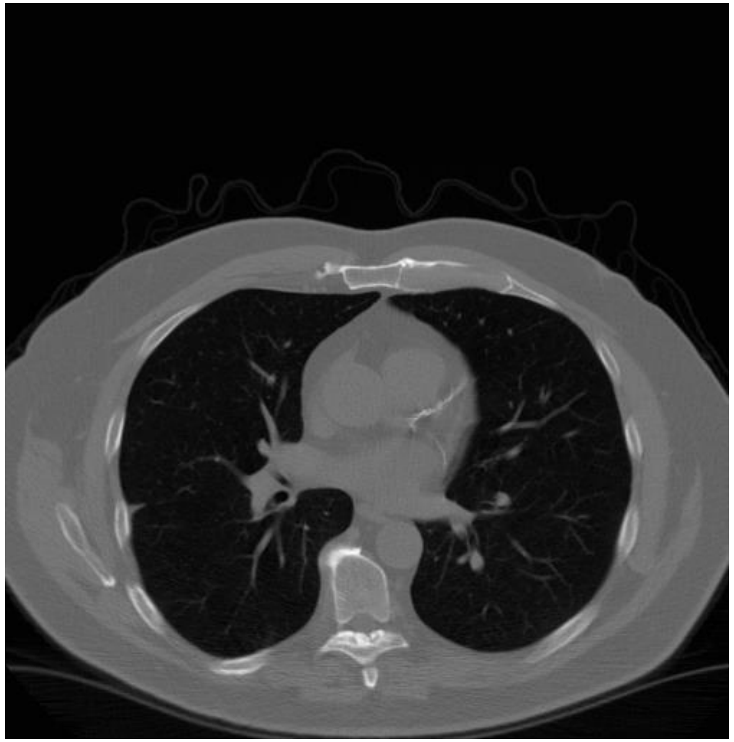
For real-time automatic cardiovascular disease (CVD) risk categorization



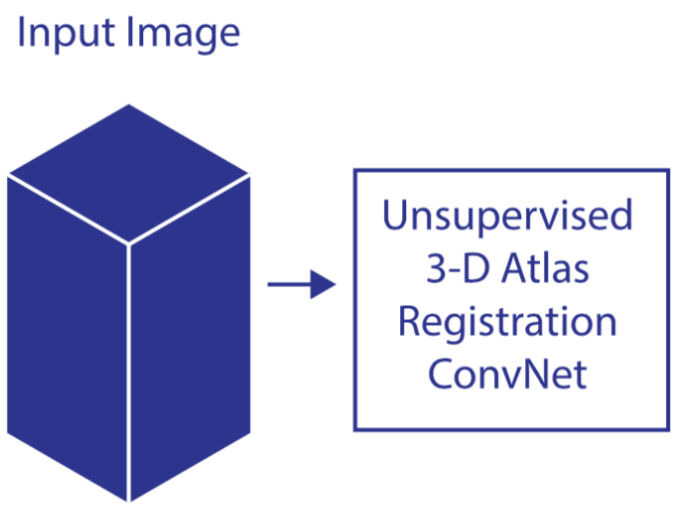
Cardiac CT

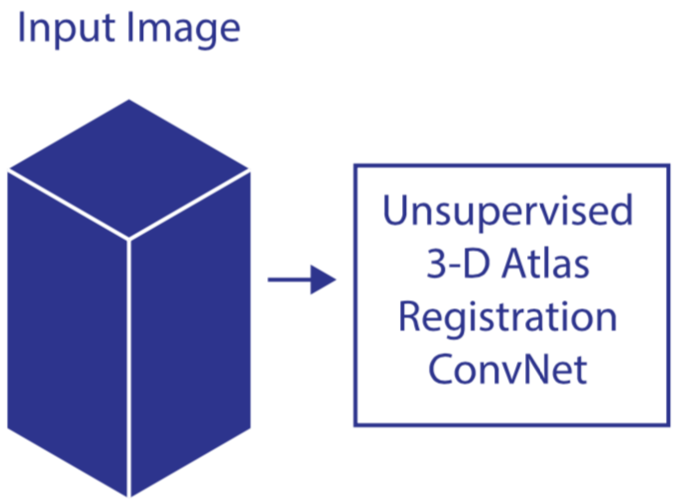


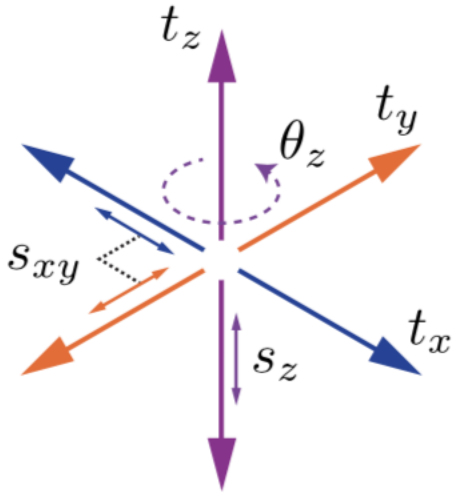
Cardiac CT

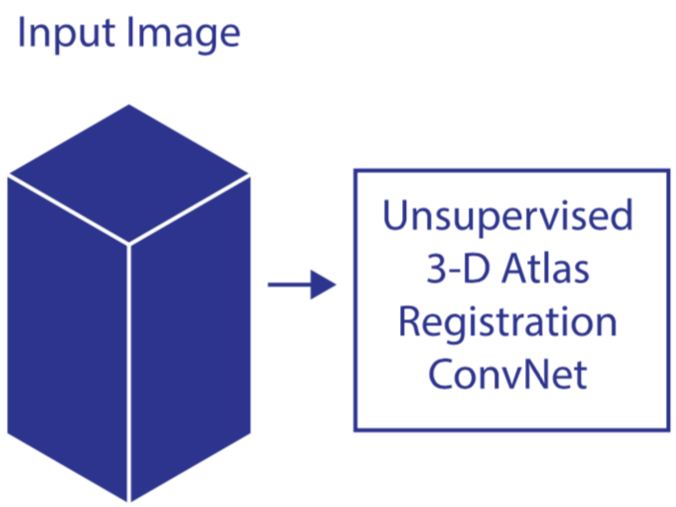


Chest CT

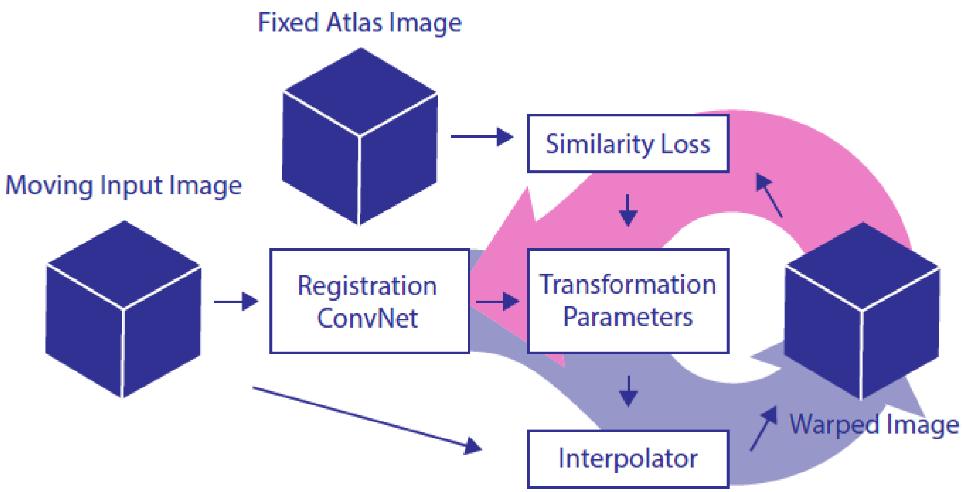

Input Image
Unsupervised
3-D Atlas
Registration
ConvNet


Input Image
Unsupervised
3-D Atlas
Registration
ConvNet

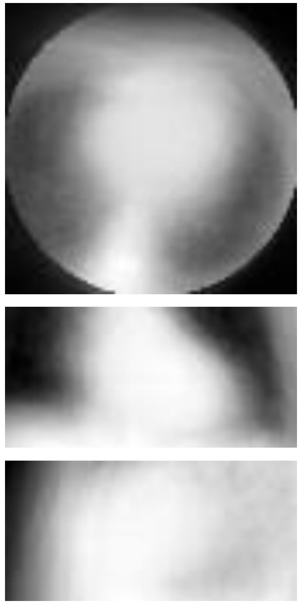

t_z
t_y
θ_z
s_{xy}
t_x
s_z


Input Image
Unsupervised
3-D Atlas
Registration
ConvNet

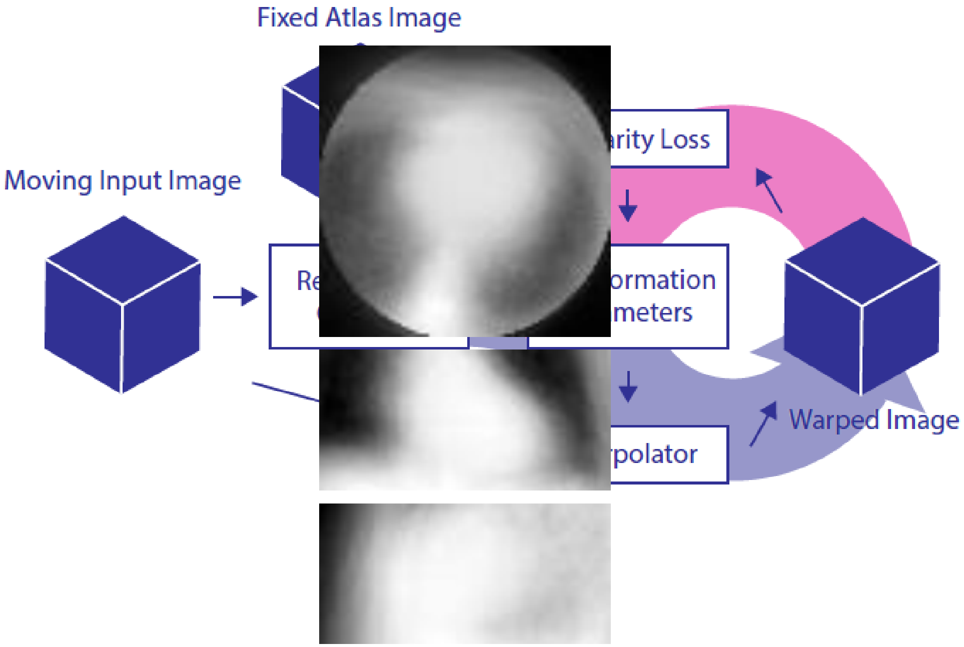
Group-wise image registration framework



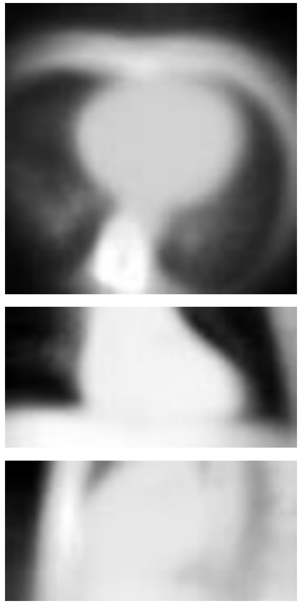
Creating an atlas-image from cardiac CT



Average of training set
before registration

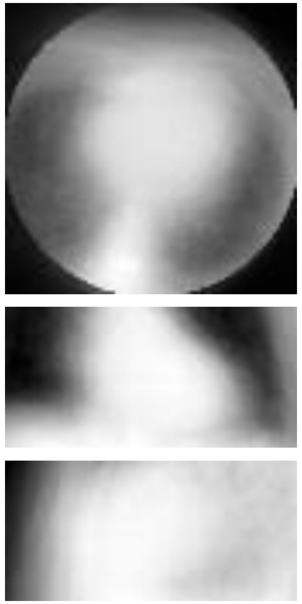


Fixed image

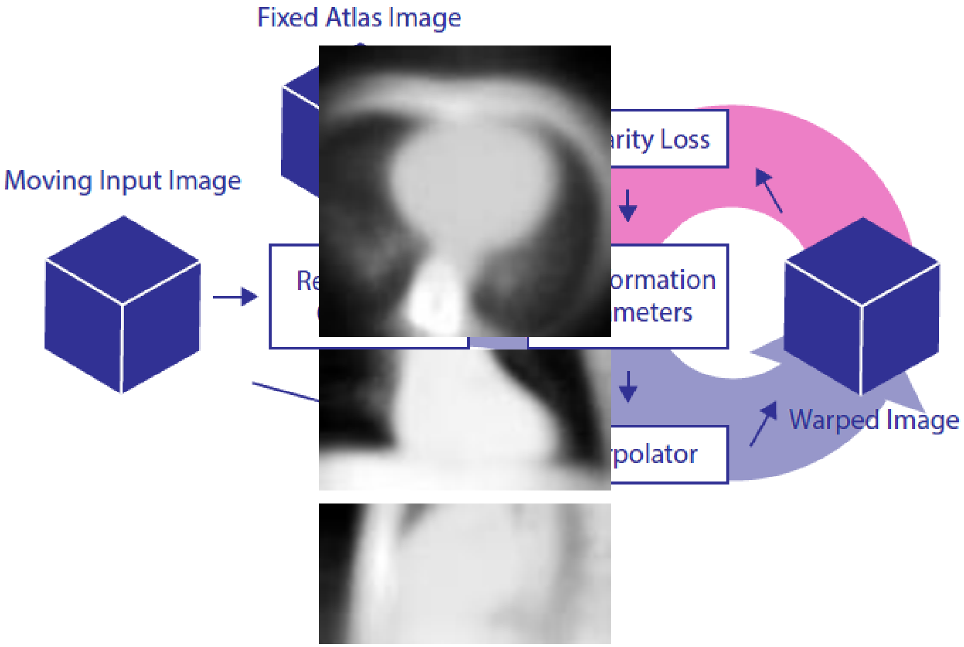


Average of training set
after registration

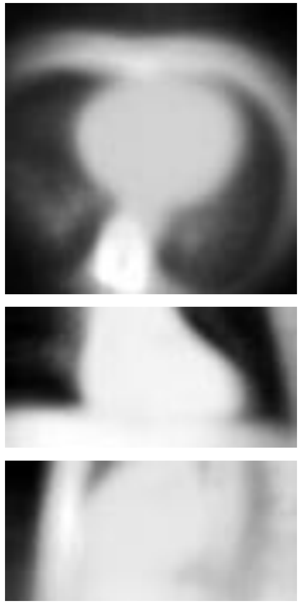
Creating an atlas-image from cardiac CT



Average of training set before registration

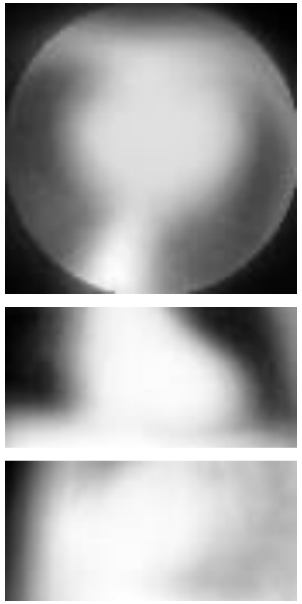


Refined atlas

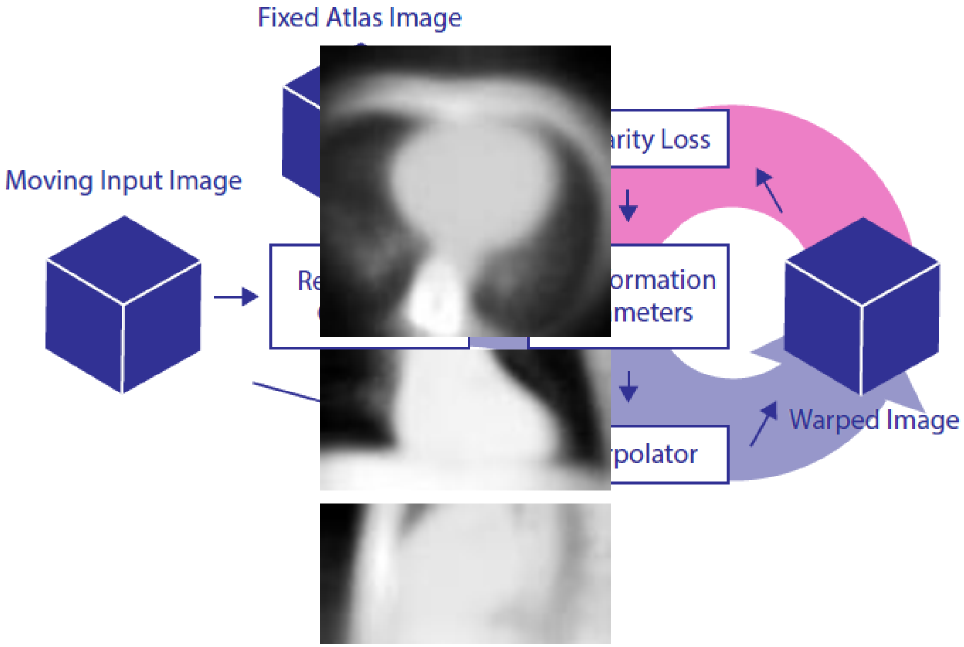


Average of training set after registration

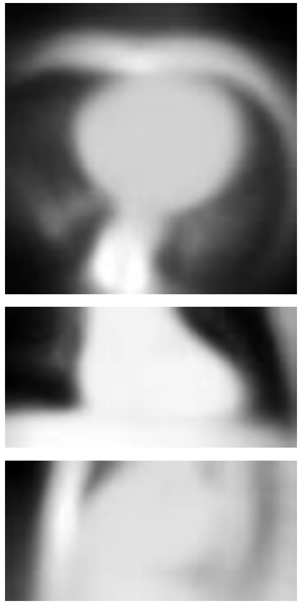
Group-wise cardiac CT registration



Average of test-set
before registration

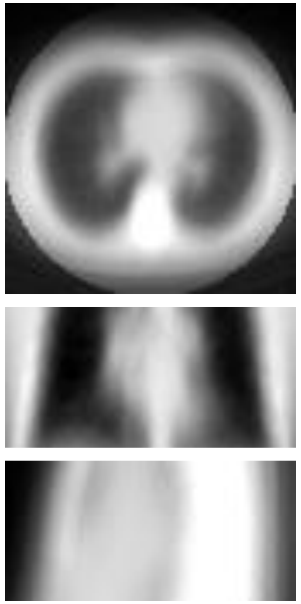


Refined atlas

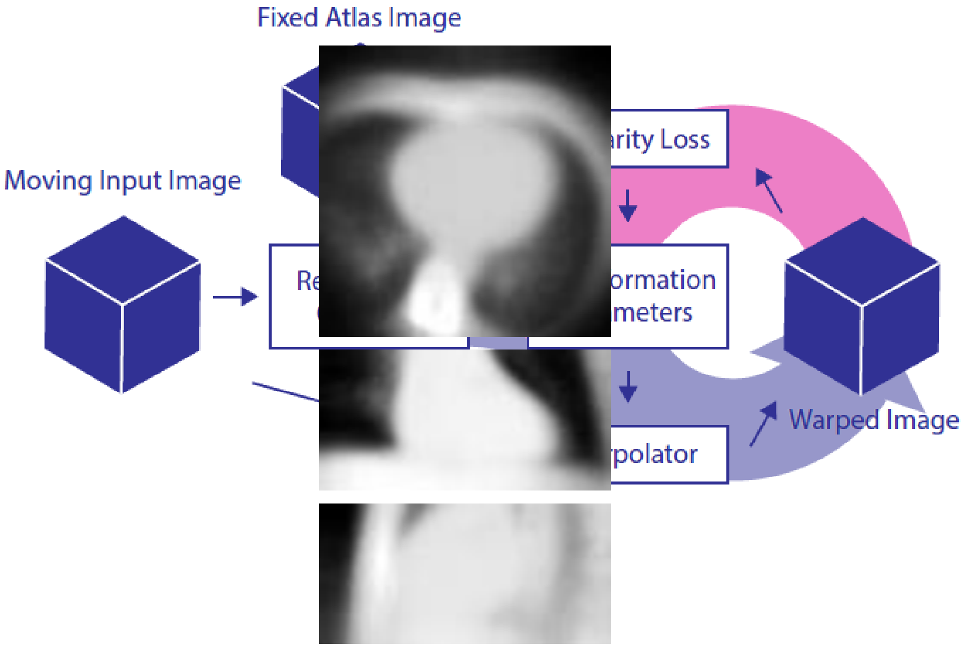


Average of test-set
after registration

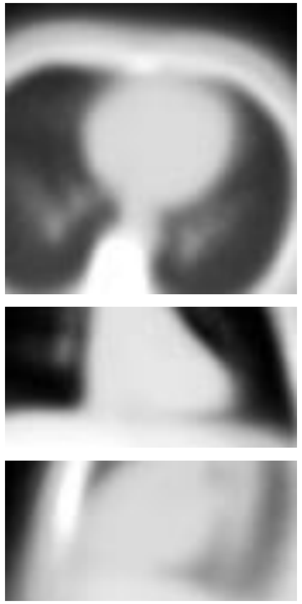
Group-wise chest CT registration



Average of moving images
before registration



Refined atlas



Average of moving images
after registration

Registration speed

- Current deep learning-based image registration is fast
  - Cardiac MR (256x256) 2D → ~10 ms
  - Cardiac MR (256x256x10) 3D single stage → ~30 ms
  - Chest CT (resampled 110^3 – 220^3) → ~0.43 s

Take-home messages

- There is not yet a one-size-fits-all solution
- Training takes some time, but inference is quick
- Inference time is mainly limited by GPU I/O-time
- High GPU-memory requirement for native resolution
- Large set of training data required, but training is unsupervised
- Easily implementable using existing building blocks from DL libraries
- Techniques from conventional image registration can be readily applied to unsupervised deep learning image registration