AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE
AGH UNIVERSITY OF SCIENCE
AND TECHNOLOGY

Deep Learning in Medical Image Analysis

2023/2024

Lecture 7: Unsupervised Learning, Medical Image Registration

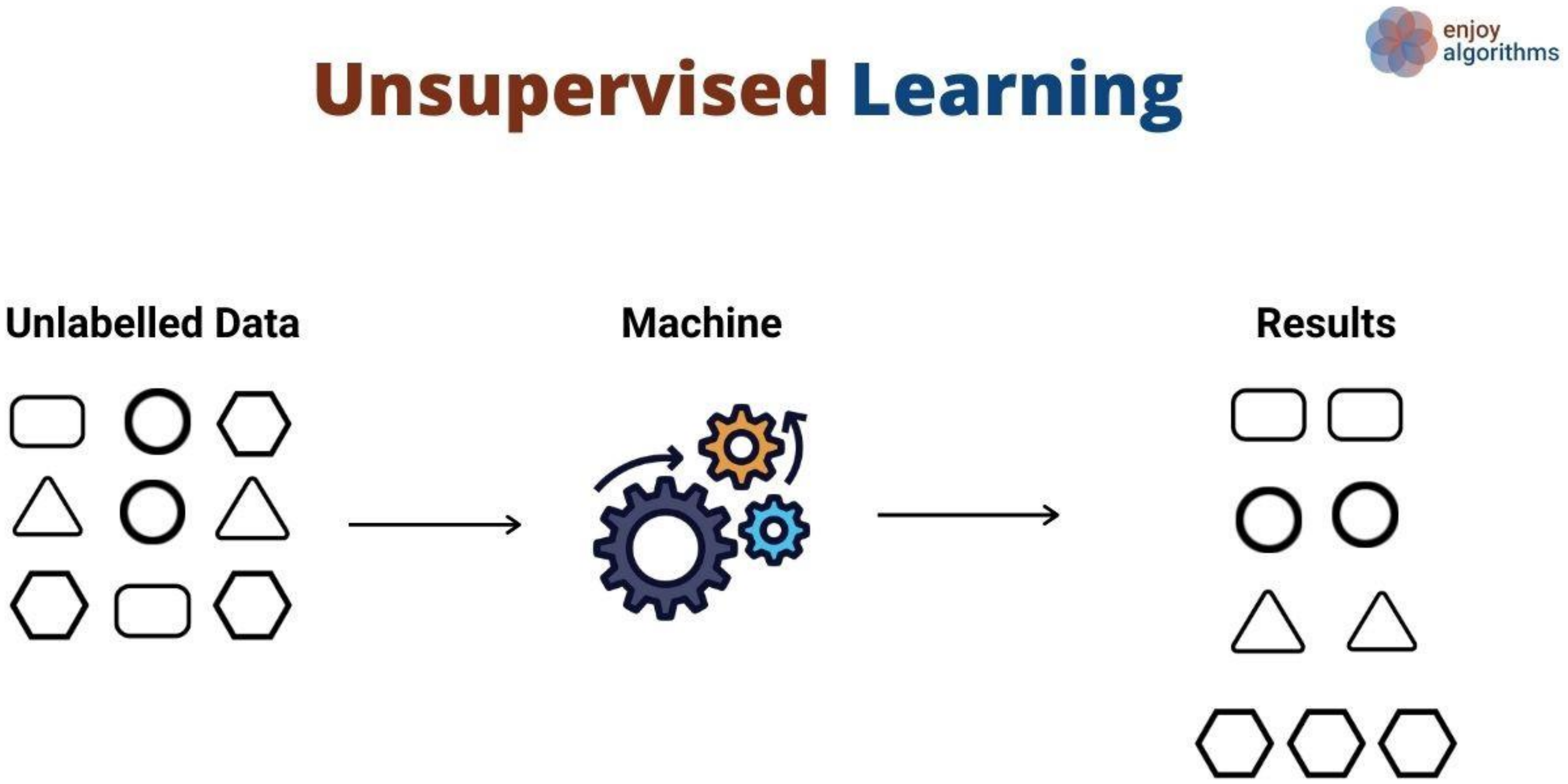
Marek Wodzinski

Goals

- Recap about the unsupervised learning
- Introduction to medical image registration
- Registration architectures
- Evaluation of the registration algorithms

Unsupervised Learning

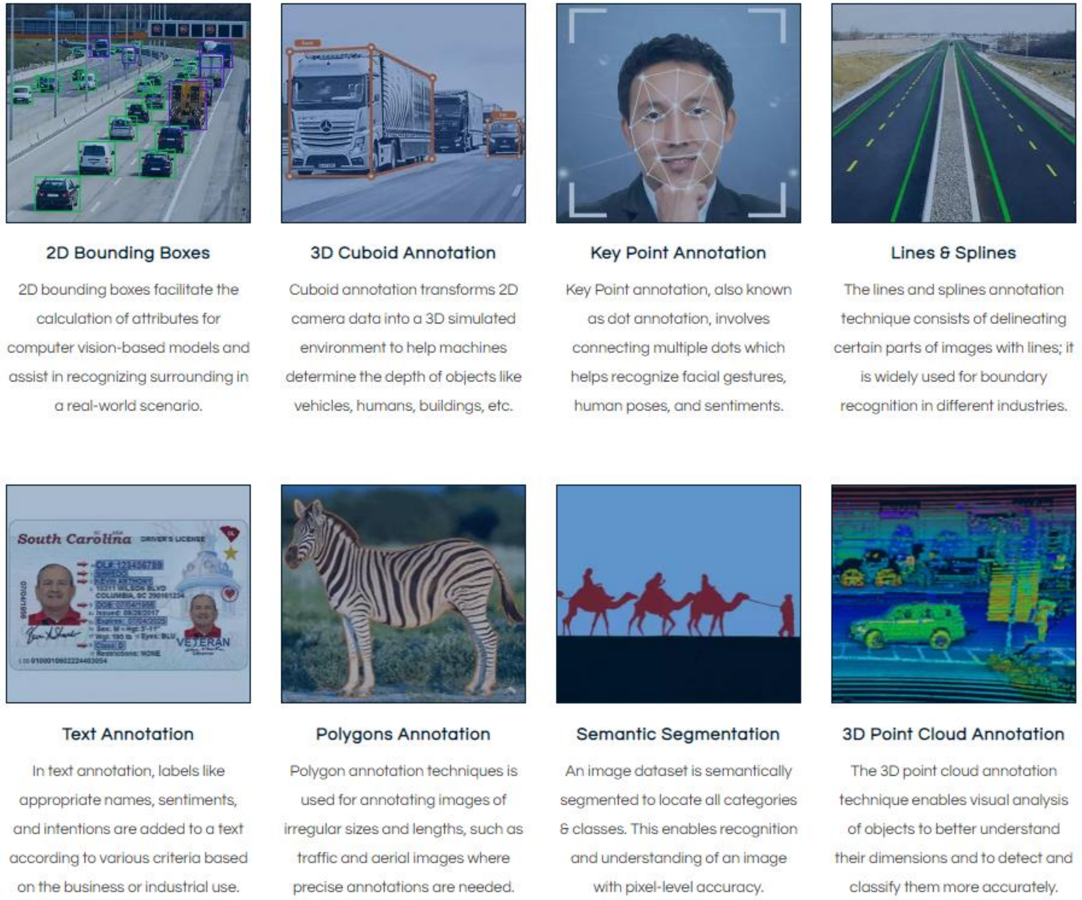
Unsupervised Learning



https://www.enjoyalgorithms.com/blogs/supervised-unsupervised-and-semisupervised-learning

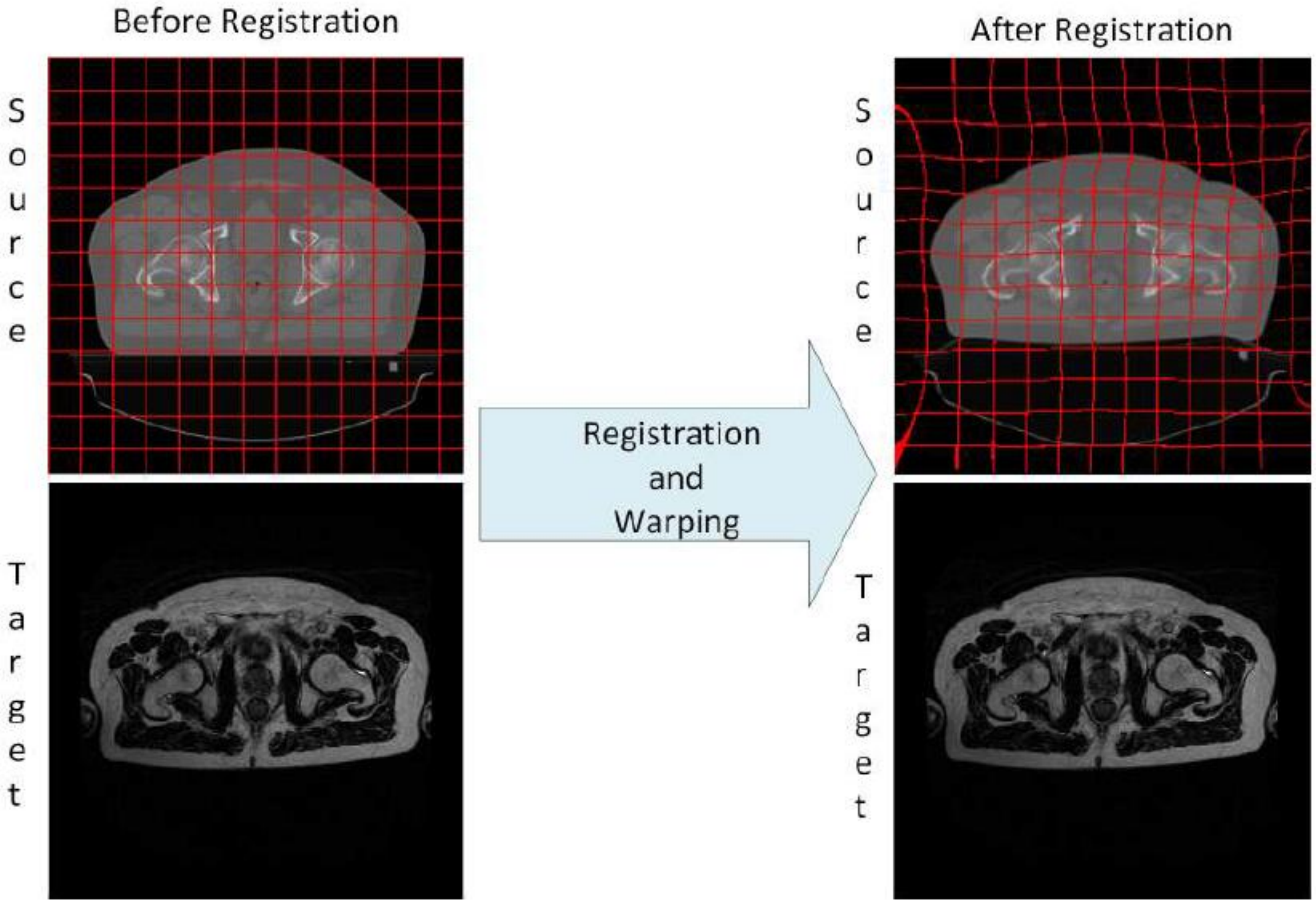
Motivation


Bounding Box
Cuboidal Annotation
Autonomous Vehicle
Image Tagging
Geospatial Annotation
Polygonal Annotation

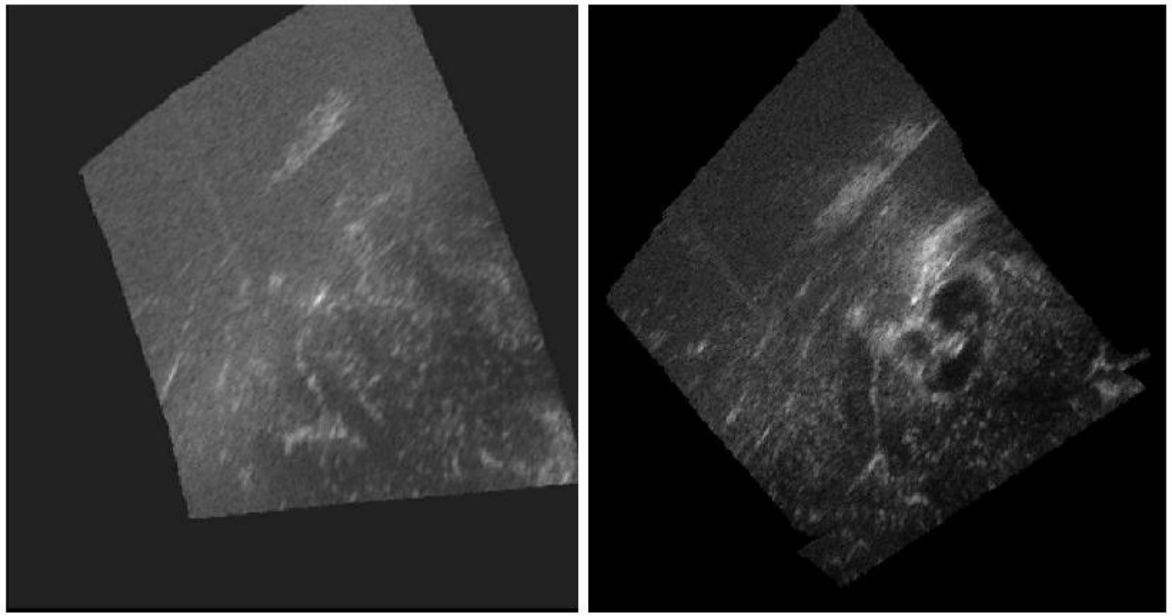

2D Bounding Boxes
2D bounding boxes facilitate the calculation of attributes for computer vision-based models and assist in recognizing surrounding in a real-world scenario.
3D Cuboid Annotation
Cuboid annotation transforms 2D camera data into a 3D simulated environment to help machines determine the depth of objects like vehicles, humans, buildings, etc.
Key Point Annotation
Key Point annotation, also known as dot annotation, involves connecting multiple dots which helps recognize facial gestures, human poses, and sentiments.
Lines & Splines
The lines and splines annotation technique consists of delineating certain parts of images with lines; it is widely used for boundary recognition in different industries.
Text Annotation
In text annotation, labels like appropriate names, sentiments, and intentions are added to a text according to various criteria based on the business or industrial use.
Polygons Annotation
Polygon annotation techniques is used for annotating images of irregular sizes and lengths, such as traffic and aerial images where precise annotations are needed.
Semantic Segmentation
An image dataset is semantically segmented to locate all categories & classes. This enables recognition and understanding of an image with pixel-level accuracy.
3D Point Cloud Annotation
The 3D point cloud annotation technique enables visual analysis of objects to better understand their dimensions and to detect and classify them more accurately.

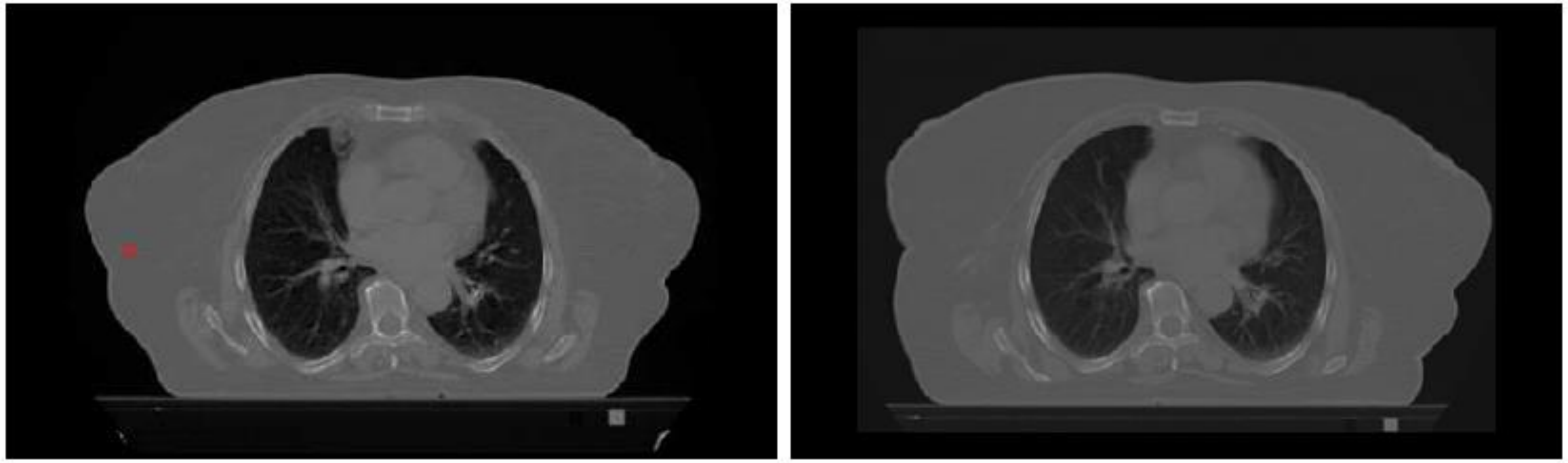
Medical Image Registration

Introduction

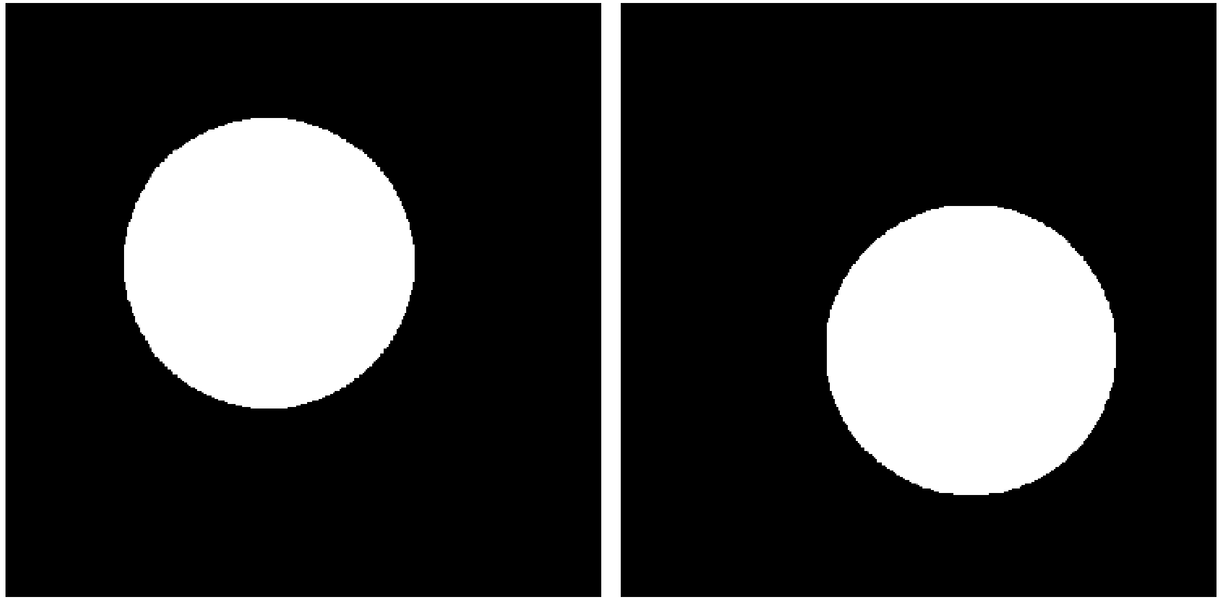


Introduction

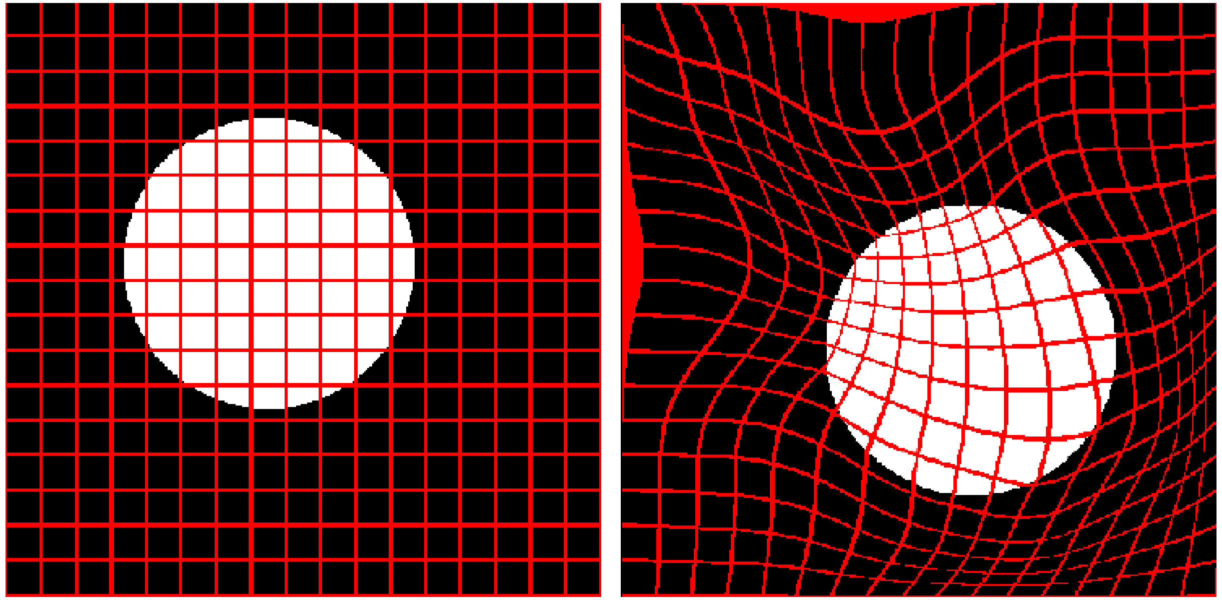




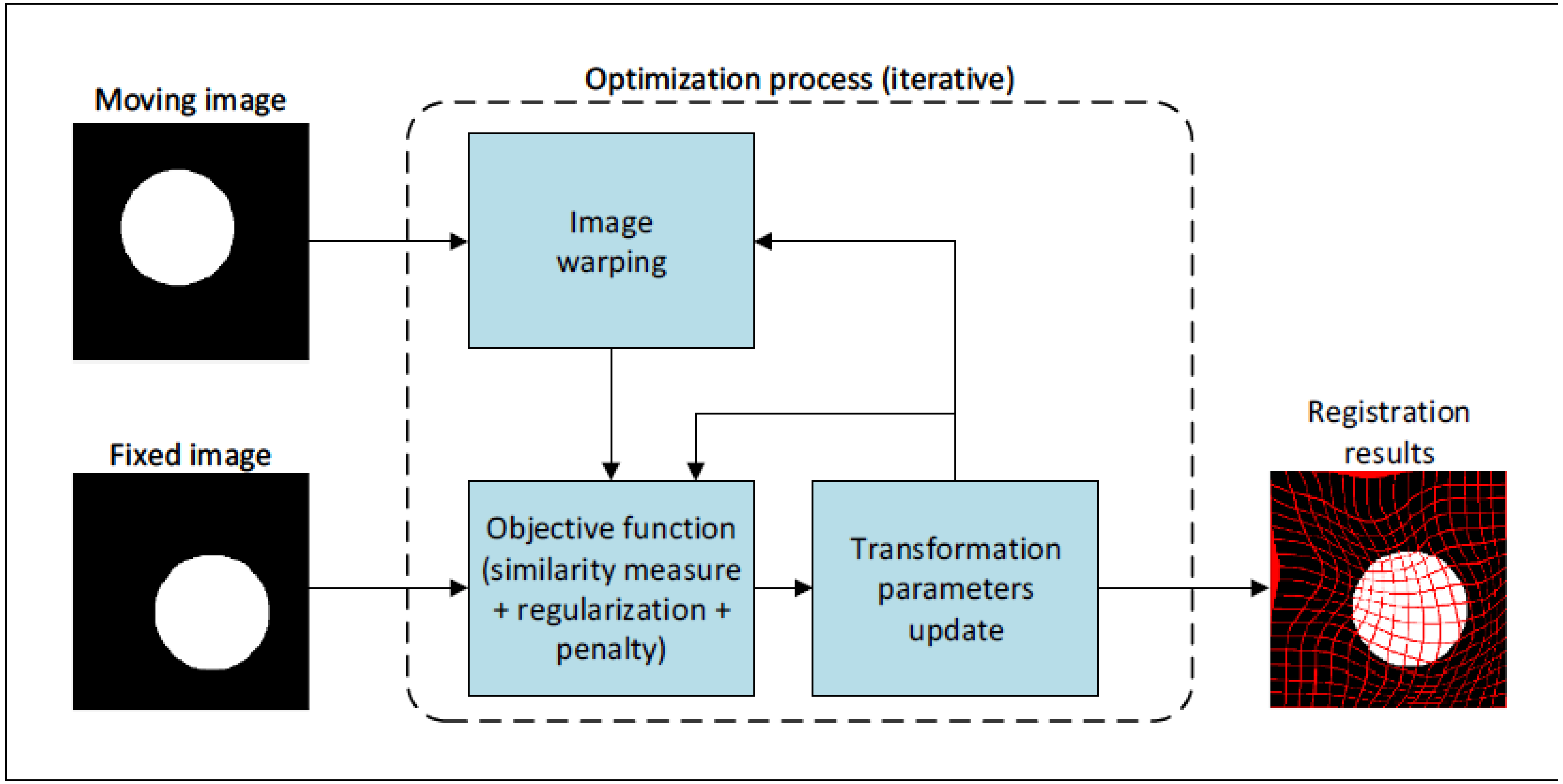
Nomenclature



(a) Moving image (b) Fixed image

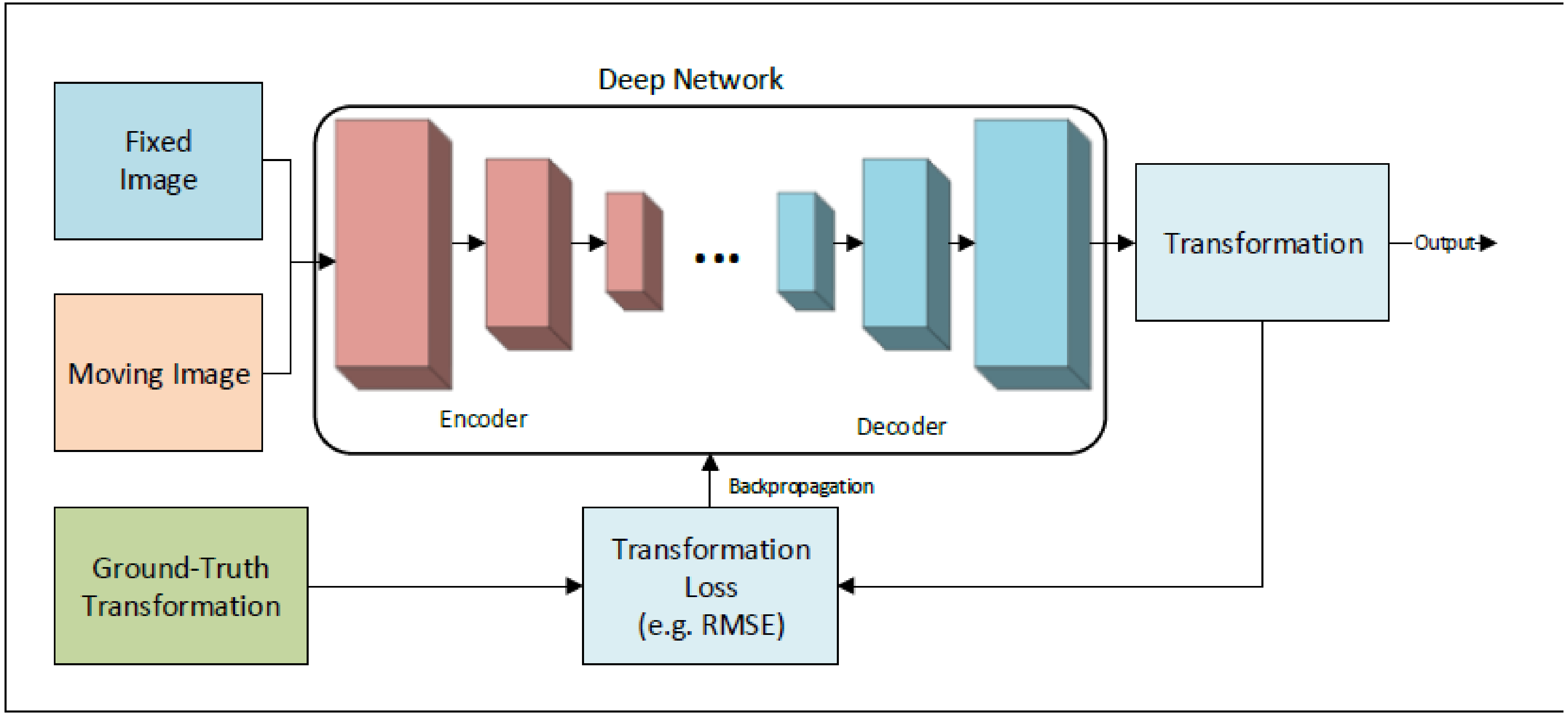


Classical Registration

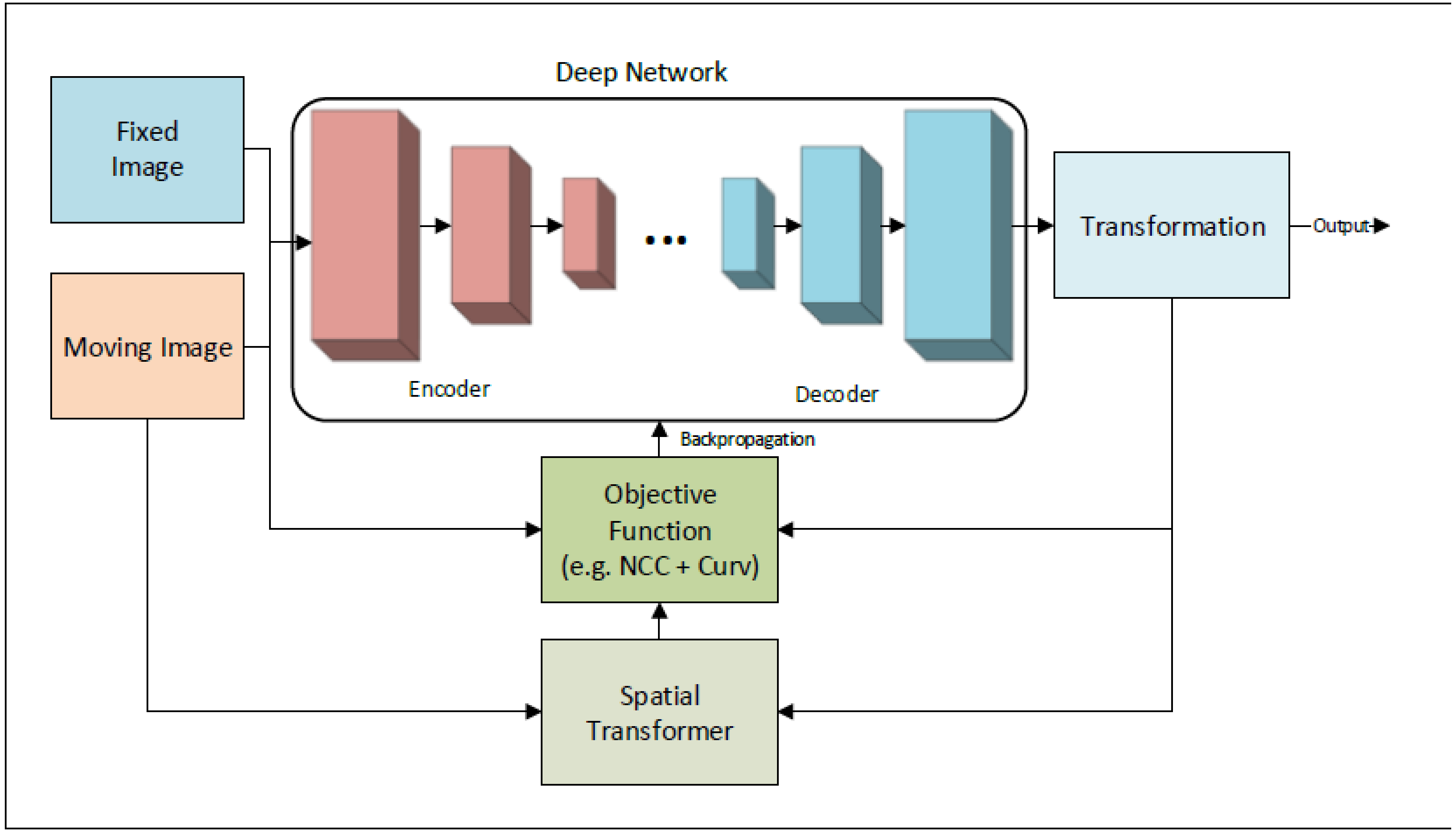


$$O(M, F, \Theta, \cdot) = C(T(M, \Theta), F) + R(\Theta) + P(\cdot) \rightarrow min,$$

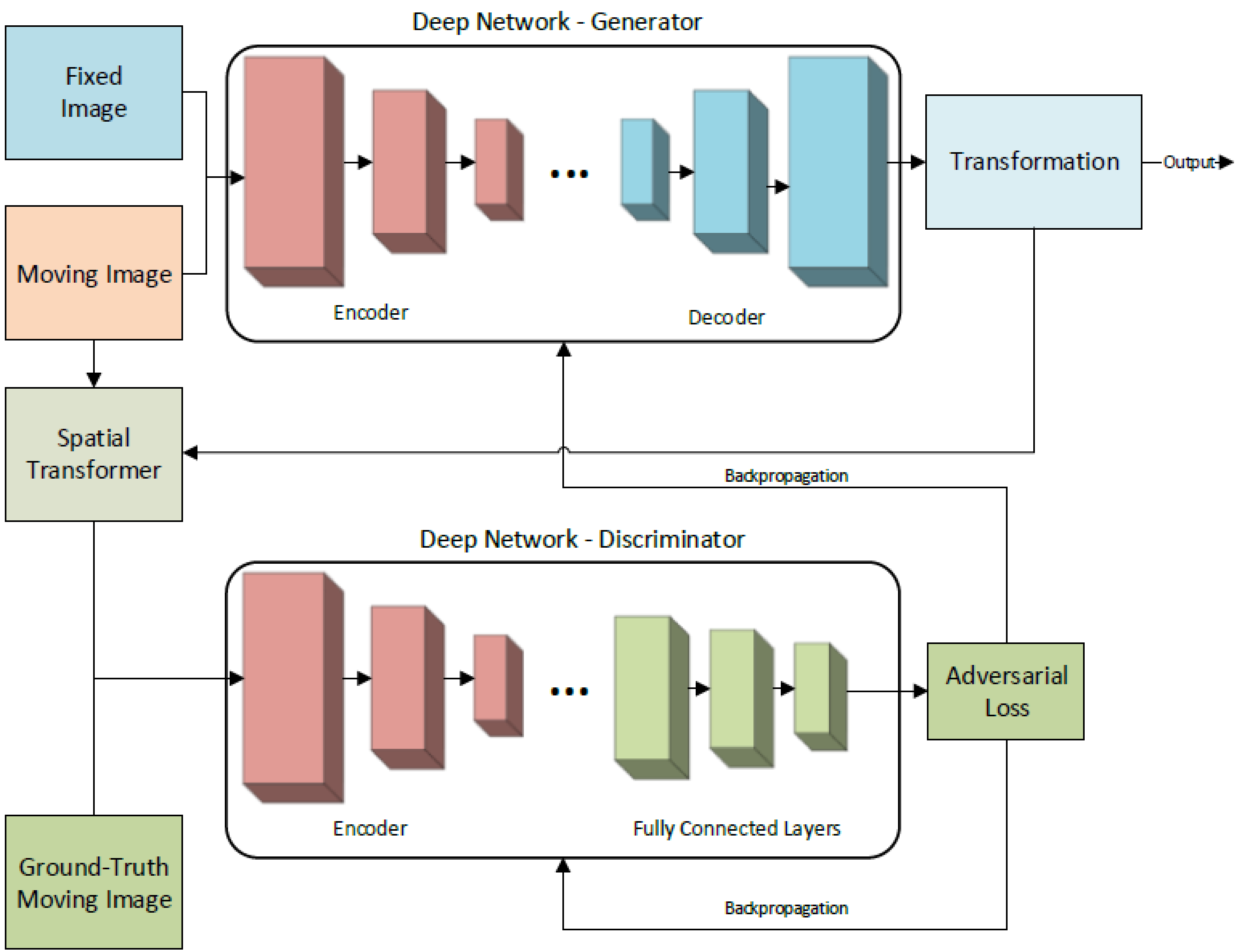
Supervised Registration



Unsupervised Registration



Adversarial Registration



Learning-based Summary

Registration Type	Similarity Measure	Ground-Truth Required	Generalizability	Training Complexity	Training Time	Memory Consumption
Supervised	✗	✓	Low	Intermediate	High	Low
Unsupervised	✓	✗	High	Low	Intermediate	Intermediate*
Adversarial	✗	✓	High	Very High	High	Very High

* The memory required for the gradient backpropagation depends partially on the similarity measure and the regularization term.

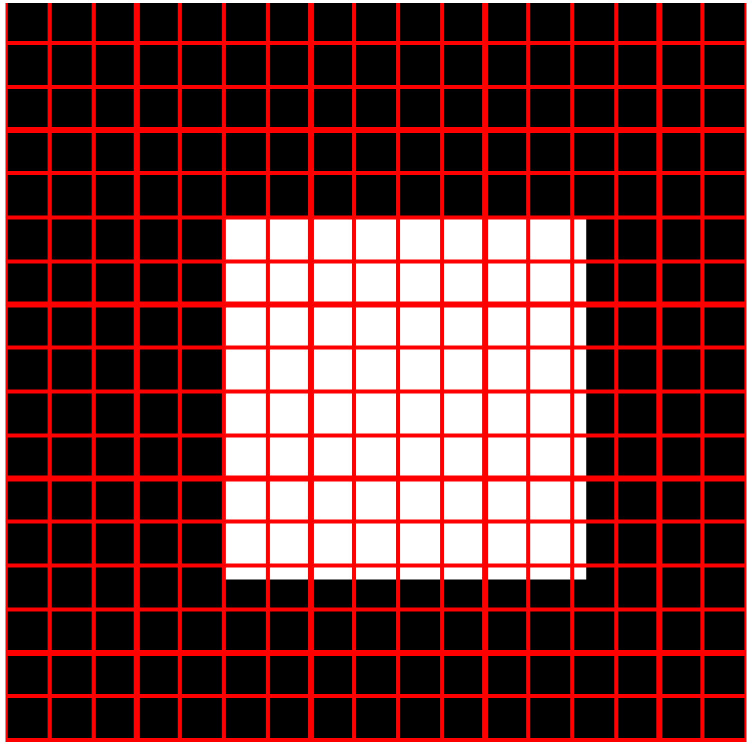
Transformation Models

Rigid Transformation

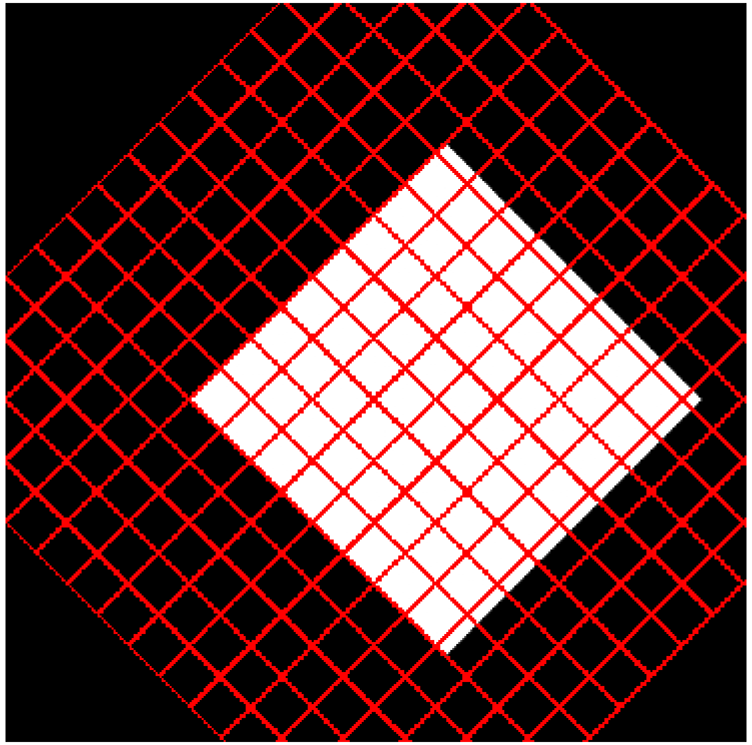
$$\mathrm{R}(\theta, \boldsymbol{t}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & t_1 \\ \sin(\theta) & \cos(\theta) & t_2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{q} = (q_1, q_2, q_3, q_4).$$

$$\mathrm{R}(\boldsymbol{q}) = \begin{bmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1 q_2 - q_3 q_4) & 2(q_1 q_3 + q_3 q_4) \\ 2(q_1 q_2 + q_3 q_4) & 1 - 2(q_1^2 + q_3^2) & 2(q_2 q_3 - q_1 q_4) \\ 2(q_1 q_3 - q_3 q_4) & 2(q_2 q_3 + q_1 q_4) & 1 - 2(q_1^2 + q_2^2) \end{bmatrix},$$



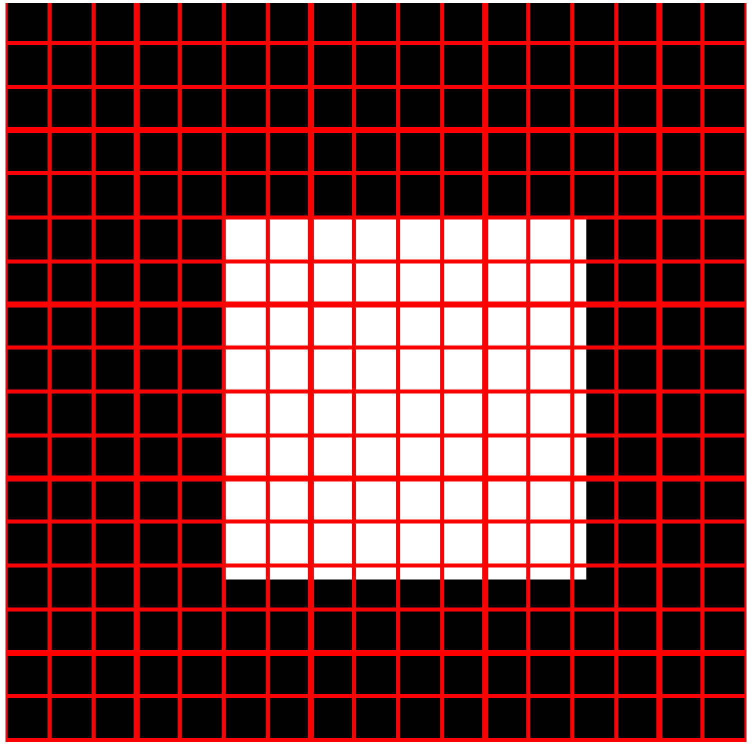
(a) Initial



(b) Rigid transform

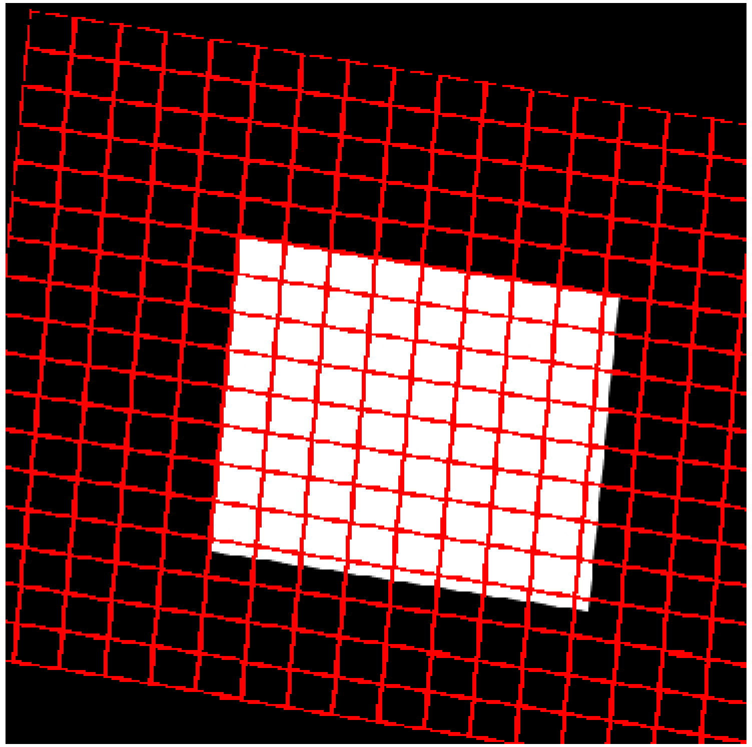
Affine Transformation

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix}$$



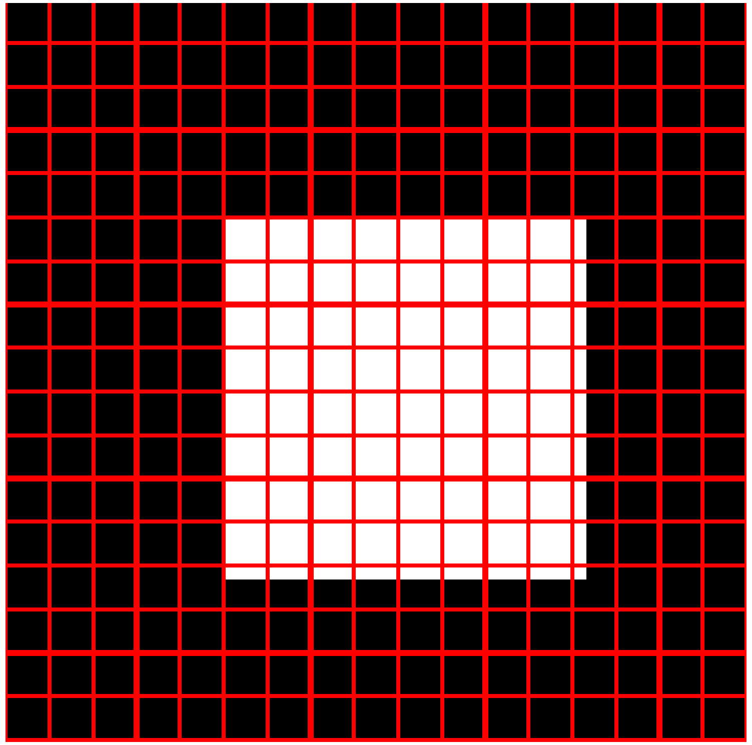
(a) Initial

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & \theta_{33} & \theta_{34} \end{bmatrix}$$

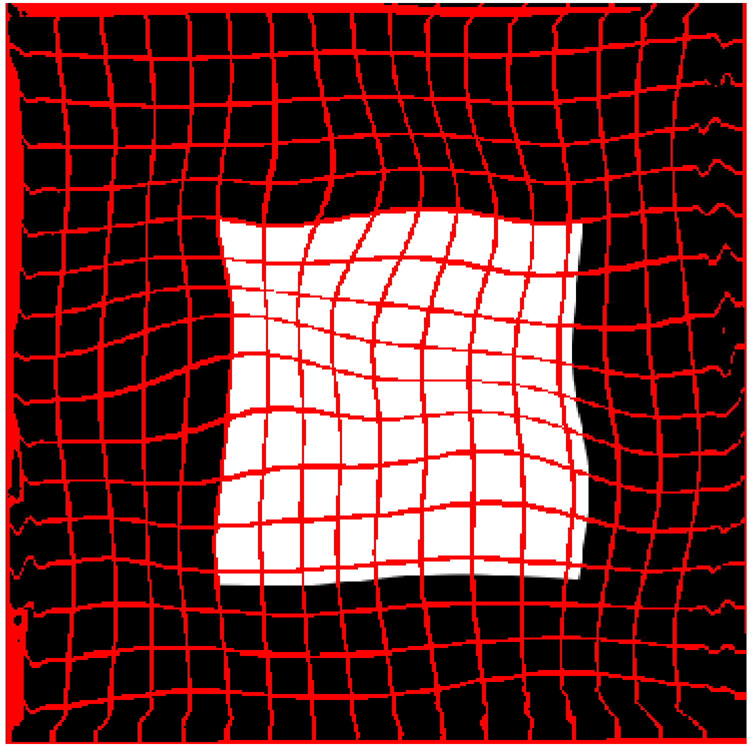


(c) Affine transform (global)

Local Affine Transformation



(a) Initial

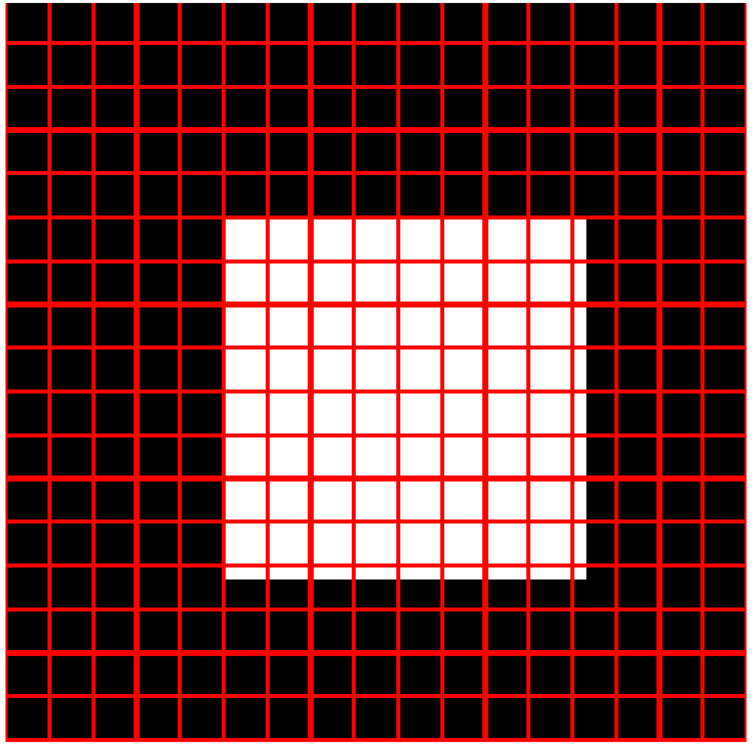


(d) Affine transform (local)

B-Splines Transformation

$$T_{ffd}(\boldsymbol{x}, \Phi) = \sum_{l=1}^{4}\sum_{m=1}^{4} B_l(u)B_m(v)\phi_{i+l,j+m},$$

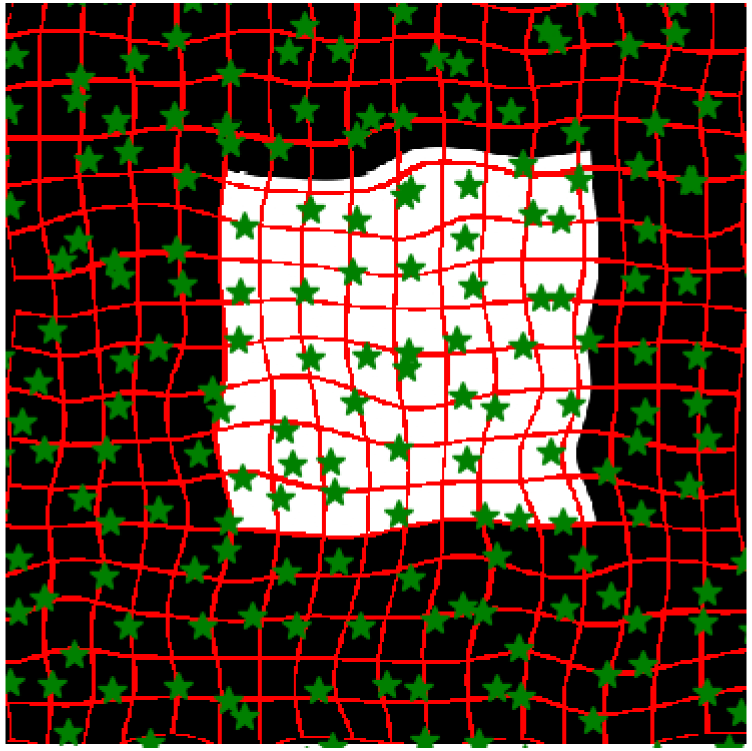
$$T_{ffd}(\boldsymbol{x}, \Phi) = \sum_{l=1}^{4}\sum_{m=1}^{4}\sum_{n=1}^{4} B_l(u)B_m(v)B_n(w)\phi_{i+l,j+m,k+n},$$



(a) Initial

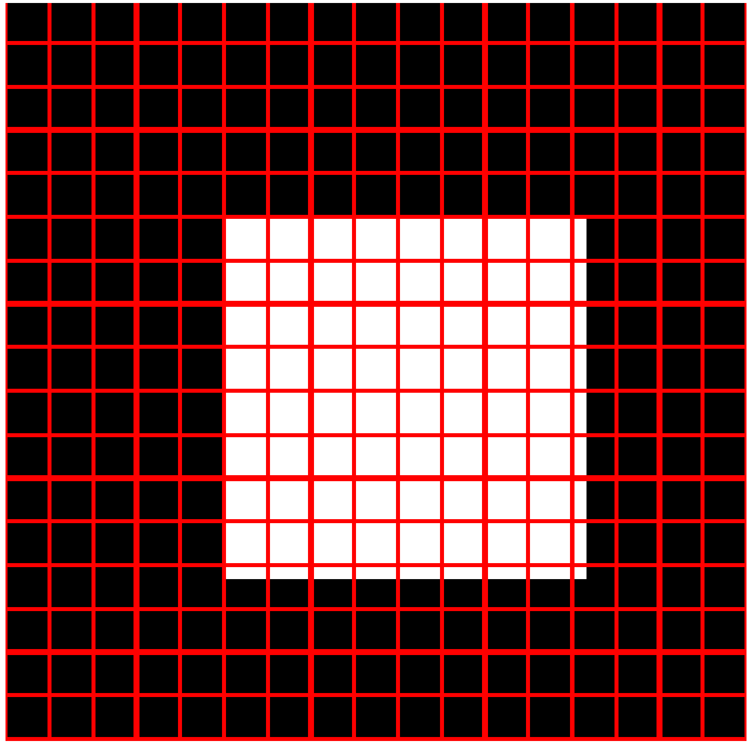
$$\begin{aligned} B_1(u) &= \frac{(1-u)^3}{6} \\ B_2(u) &= \frac{\left(3u^3-6u^2+4\right)}{6} \\ B_3(u) &= \frac{\left(-3u^3+3u^2+3u+1\right)}{6} \\ B_4(u) &= \frac{u^3}{6} \end{aligned}$$

$$\begin{aligned} \frac{\partial T_{ffd}(x;\Phi)}{\partial \phi_1(i,j,k)} &= B_p(u)B_q(v)B_r(w) \quad \text{for} \quad x,y,z \in \text{ROI} \\ \frac{\partial T_{ffd}(x;\Phi)}{\partial \phi_2(i,j,k)} &= B_p(u)B_q(v)B_r(w) \quad \text{for} \quad x,y,z \in \text{ROI} \\ \frac{\partial T_{ffd}(x;\Phi)}{\partial \phi_3(i,j,k)} &= B_p(u)B_q(v)B_r(w) \quad \text{for} \quad x,y,z \in \text{ROI} \end{aligned}$$



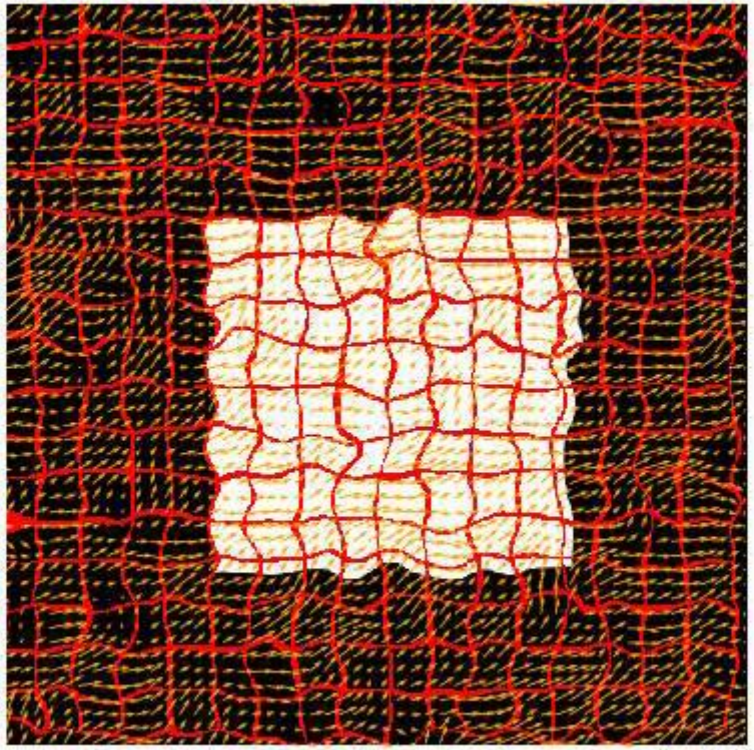
(e) Cubic B-Splines

Dense Transformation



(a) Initial

$$y(\boldsymbol{x}) = \boldsymbol{x} + u(\boldsymbol{x})$$
$$\boldsymbol{x}' = y(\boldsymbol{x}),$$



(f) Dense deformation

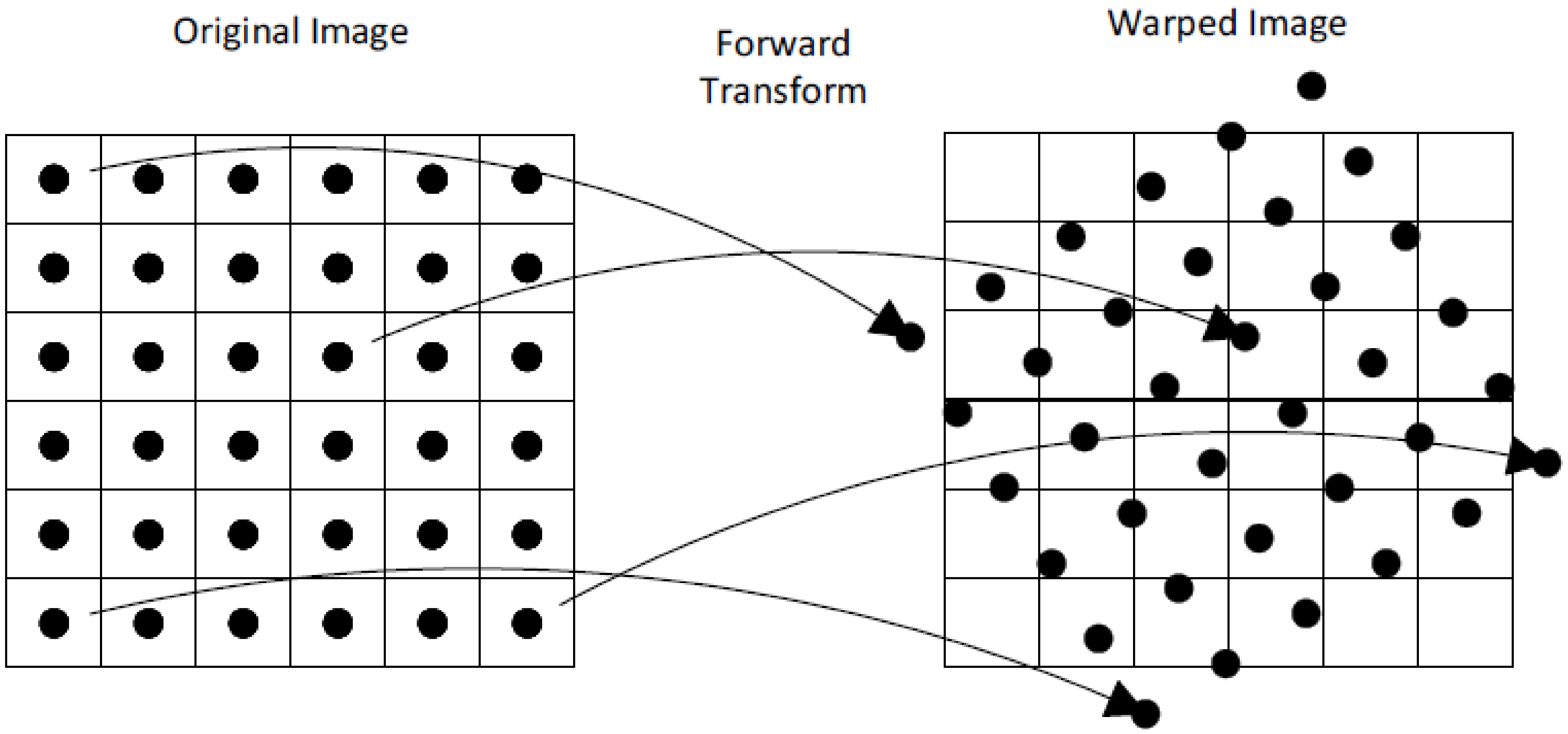
Transformation Models

Transformation	Equation	No. Parameters 2-D (3-D)	Computational Complexity	Application
Rigid	2.3	3 (6)	Low	Initial alignment
Affine - global	2.9	6 (12)	Low	Initial alignment
Affine - local	2.9	6 (12) * no. patches	Intermediate/High*	Deformable registration
B-Splines FFD	2.12	2 (3) * no. control points	Low/Intermediate**	Deformable registration
Dense deformation field	2.15	2 (3) * no. pixels/voxels	High	Deformable registration

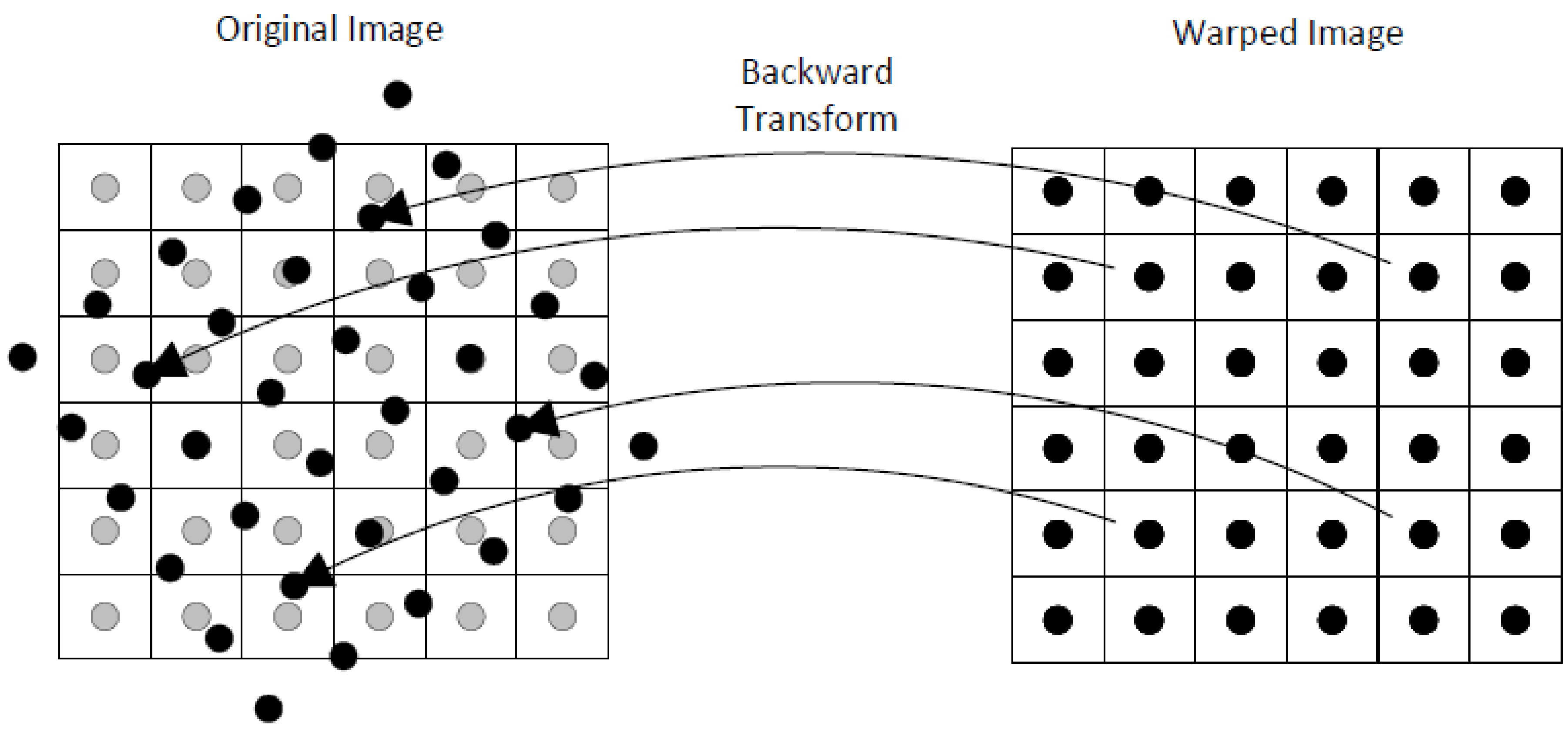
* Depends on the patch size.

** Depends on the splines order and the spacing between control points.

Forward Transformation



Backward Transformation



Cost Functions

(di)Similarity Metrics

Similarity Measure	Equation	Modalities	Computational Complexity	Application
SSD (MSE) [Szeliski, 2010]	2.17	Unimodal	Very Low	Simple registration without intensity variations
SAD [Szeliski, 2010]	2.18	Unimodal	Very Low	Same as SSD
NCC [Szeliski, 2010]	2.19	Unimodal*	Low	Any application with a linear intensity relation between corresponding structures
MI [Viola and Wells III, 1997]	2.20	Multimodal	High	General multimodal problems
NMI [Studholme et al., 1999]	2.21	Multimodal	Very High	Same as MI + more resistant to non-overlapping region of interest
NGF [Haber and Modersitzki, 2006]	2.22	Multimodal	Low	Images with the intensity changes occurring at the corresponding locations
MIND-SSD [Heinrich et al., 2012]	2.25	Multimodal	Intermediate	General multimodal problems, more accurate and faster than MI/NMI in numerous applications

* Can be used for multimodal problems after applying an intensity transformation.

(di)Similarity Metrics

$$MSE(M(\boldsymbol{x}), F(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N}(M(\boldsymbol{x}_i) - F(\boldsymbol{x}_i))^2,$$

$$SAD(M(\boldsymbol{x}), F(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N}|M(\boldsymbol{x}_i) - F(\boldsymbol{x}_i)|.$$

(di)Similarity Metrics

$$NCC(M(\boldsymbol{x}), F(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N} \frac{(M(\boldsymbol{x}_i) - \mu_M)(F(\boldsymbol{x}_i) - \mu_F)}{\sigma_M \sigma_F},$$

(di)Similarity Metrics

$$MI(M(\boldsymbol{x}), F(\boldsymbol{x})) = H(p(M(\boldsymbol{x}))) + H(p(F(\boldsymbol{x}))) - H(p(M(\boldsymbol{x}), F(\boldsymbol{x}))),$$

$$NMI(M(\boldsymbol{x}), F(\boldsymbol{x})) = \frac{H(p(M(\boldsymbol{x}))) + H(p(F(\boldsymbol{x})))}{H(p(M(\boldsymbol{x}), F(\boldsymbol{x})))}.$$

(di)Similarity Metrics

$$NGF(M(\boldsymbol{x}), F(\boldsymbol{x})) = \sum_{i=1}^{N} 1 - \left(\frac{\langle \nabla M(\boldsymbol{x}_i), \nabla F(\boldsymbol{x}_i) \rangle}{\| \nabla M(\boldsymbol{x}_i) \|_\epsilon \| \nabla F(\boldsymbol{x}_i) \|_\epsilon} \right)^2,$$

(di)Similarity Metrics

$$MIND(I(\boldsymbol{x}), \boldsymbol{x}, \boldsymbol{R}) = \frac{1}{n}\exp(-\frac{D_p(I(\boldsymbol{x}), \boldsymbol{x}, \boldsymbol{R})}{V(I(\boldsymbol{x}))},$$

$$D_p(I(\boldsymbol{x}), \boldsymbol{x}, \boldsymbol{r}) = \sum_{\boldsymbol{r} \in \boldsymbol{R}} (I(\boldsymbol{x}) - I(\boldsymbol{x} + \boldsymbol{r}))^2.$$

$$MIND_{SSD}(M(\boldsymbol{x}), F(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N}(MIND((M(\boldsymbol{x}), \boldsymbol{x}_i, \boldsymbol{R}) - MIND(F(\boldsymbol{x}), \boldsymbol{x}_i, \boldsymbol{R}))^2,$$

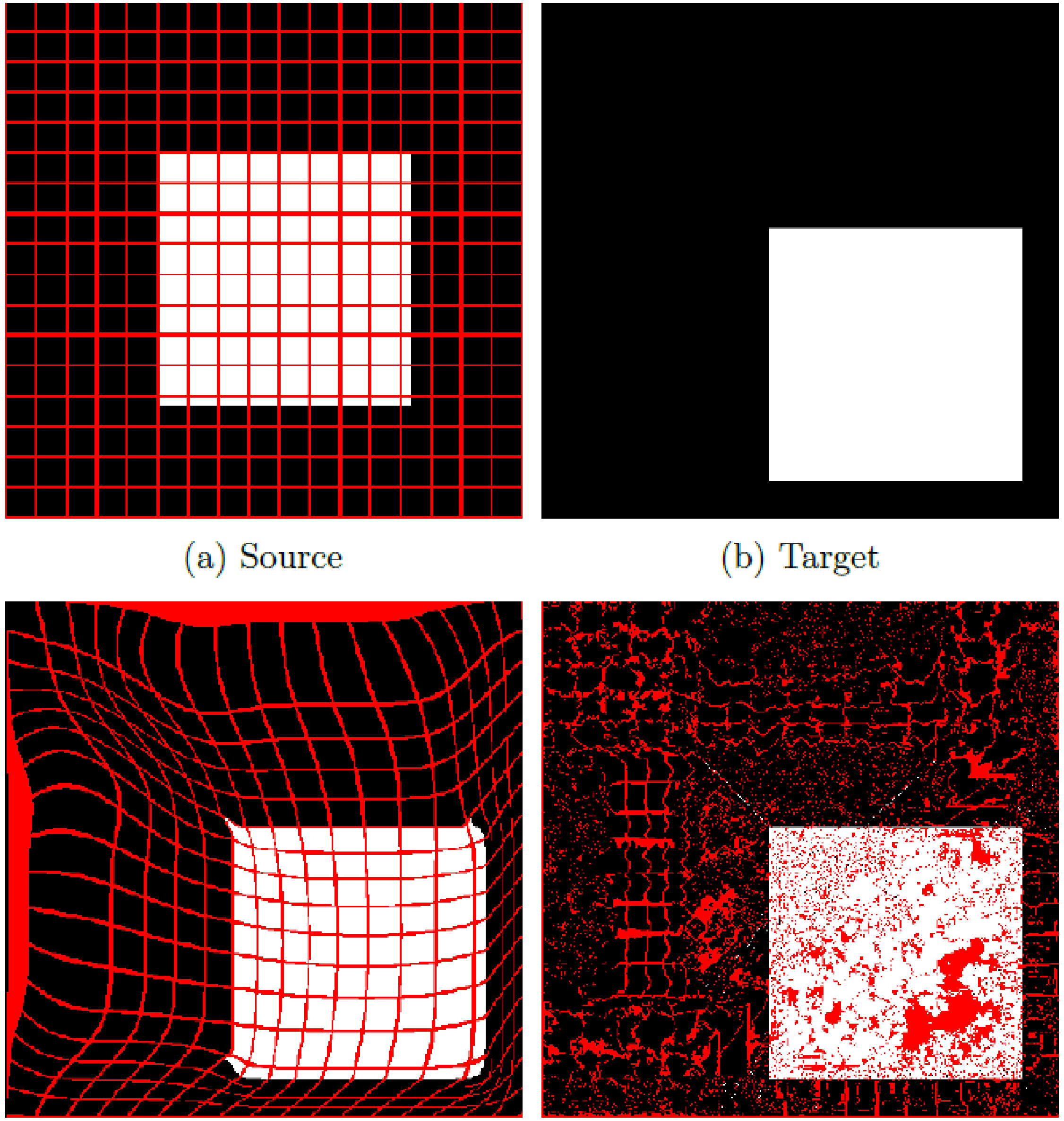
(di)Similarity Metrics

Similarity Measure	Equation	Modalities	Computational Complexity	Application
SSD (MSE) [Szeliski, 2010]	2.17	Unimodal	Very Low	Simple registration without intensity variations
SAD [Szeliski, 2010]	2.18	Unimodal	Very Low	Same as SSD
NCC [Szeliski, 2010]	2.19	Unimodal*	Low	Any application with a linear intensity relation between corresponding structures
MI [Viola and Wells III, 1997]	2.20	Multimodal	High	General multimodal problems
NMI [Studholme et al., 1999]	2.21	Multimodal	Very High	Same as MI + more resistant to non-overlapping region of interest
NGF [Haber and Modersitzki, 2006]	2.22	Multimodal	Low	Images with the intensity changes occurring at the corresponding locations
MIND-SSD [Heinrich et al., 2012]	2.25	Multimodal	Intermediate	General multimodal problems, more accurate and faster than MI/NMI in numerous applications

* Can be used for multimodal problems after applying an intensity transformation.

Regularization

Motivation



(a) Source (b) Target

(c) Registered (regularization) (d) Registered (no regularization)

Regularization Functions

- Diffusion

$$R_{diff}(y(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{D}\sum_{k=1}^{D}\alpha_{diff}\left(\frac{\delta y(\boldsymbol{x}_{ij})}{\delta x_k}\right)^2,$$

- Curvature

$$R_{curv}(y(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{D}\sum_{k=1}^{D}\sum_{l=1}^{D}\alpha_{curv}\left(\frac{\delta^2 y(\boldsymbol{x}_{ij})}{\delta x_k \delta x_l}\right)^2,$$

- Elastic

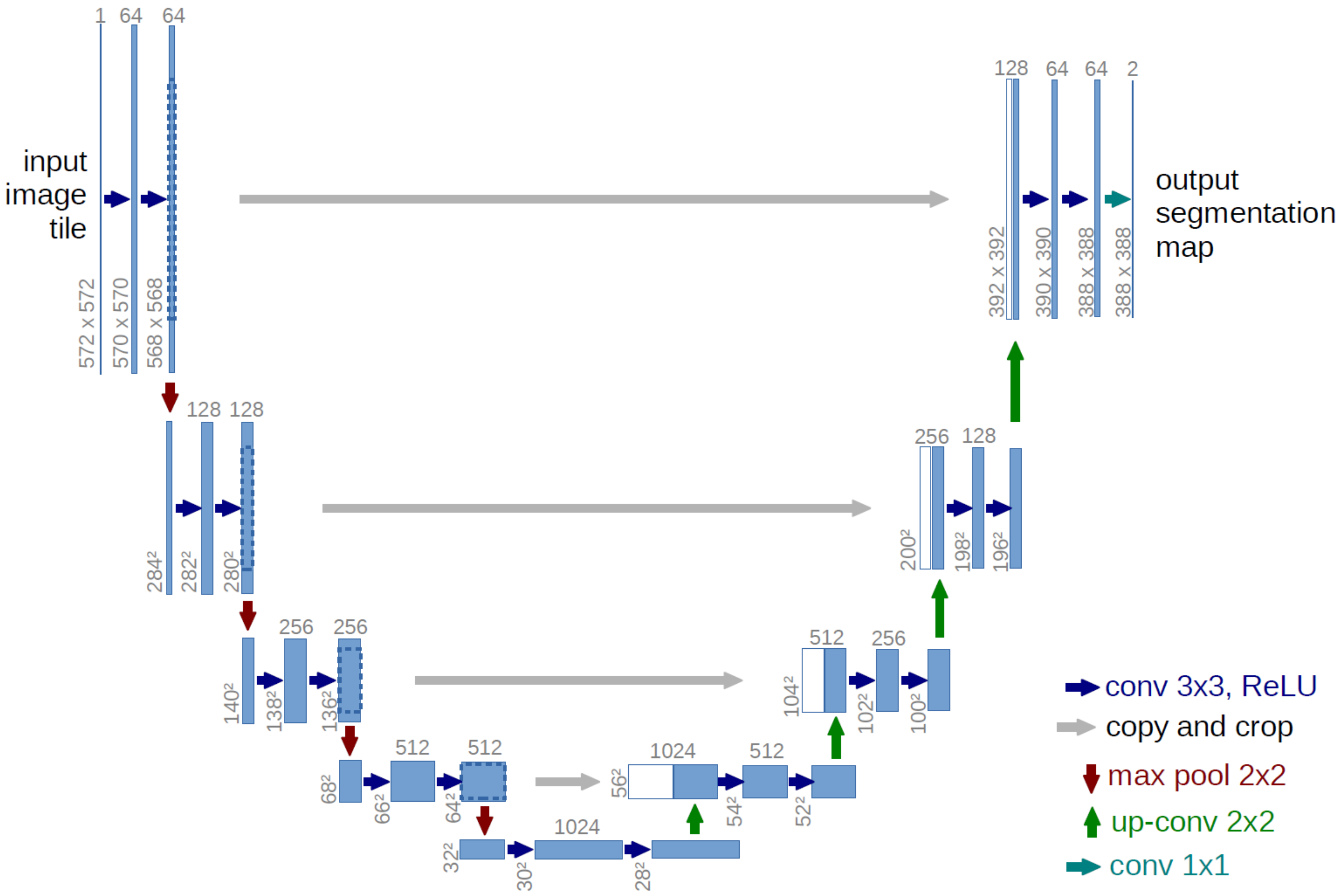
$$R_{elas}(y(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{D}\sum_{k=1}^{D}\left(\alpha_{elas}\left(\frac{\delta y(\boldsymbol{x}_{ij})}{\delta x_j}\right)\left(\frac{\delta y(\boldsymbol{x}_{ik})}{\delta x_k}\right) + \frac{\beta_{elas}}{2}\left(\frac{\delta y(\boldsymbol{x}_{ij})}{\delta x_k} + \frac{\delta y(\boldsymbol{x}_{ik})}{\delta x_j}\right)^2\right),$$

- Implicit (e.g. Gaussian)

$$y(\boldsymbol{x}) = y(\boldsymbol{x}) * G(\boldsymbol{x}, \sigma),$$

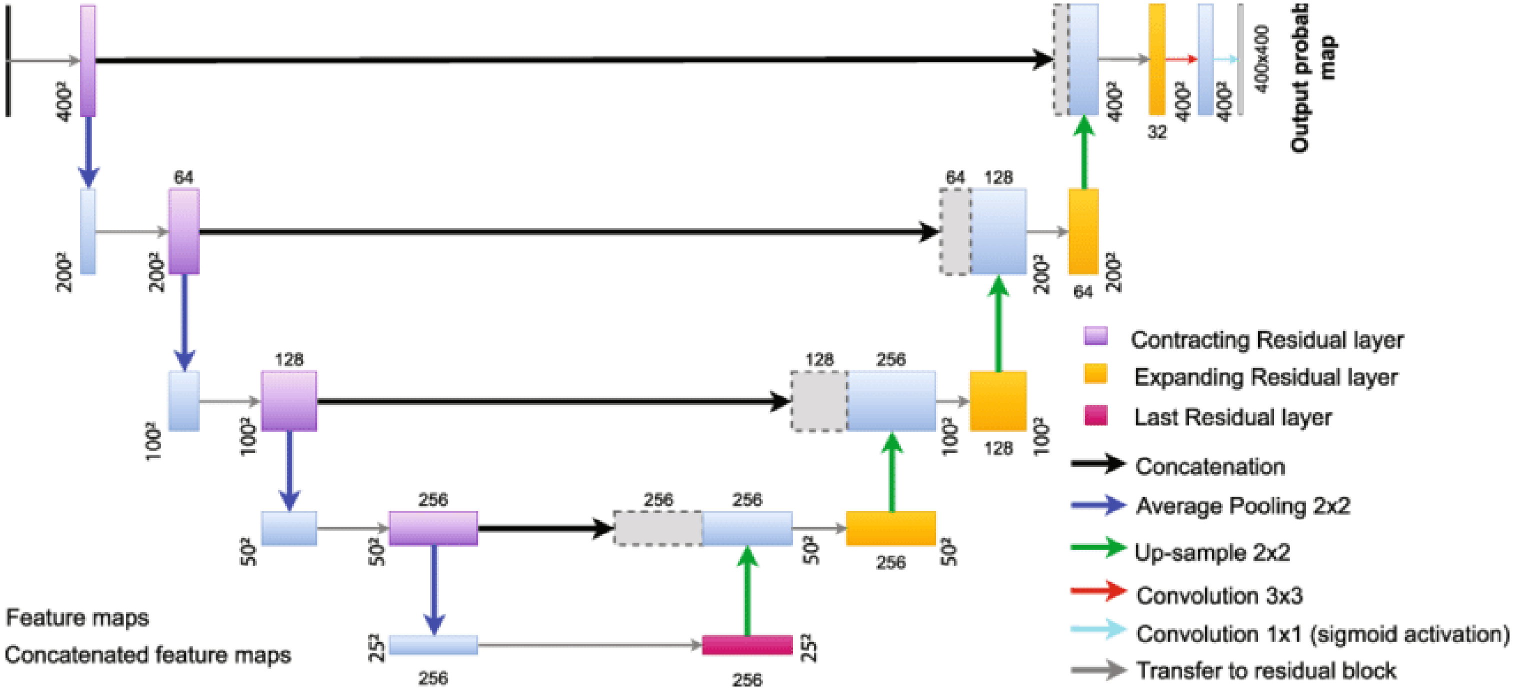
Registration Architectures

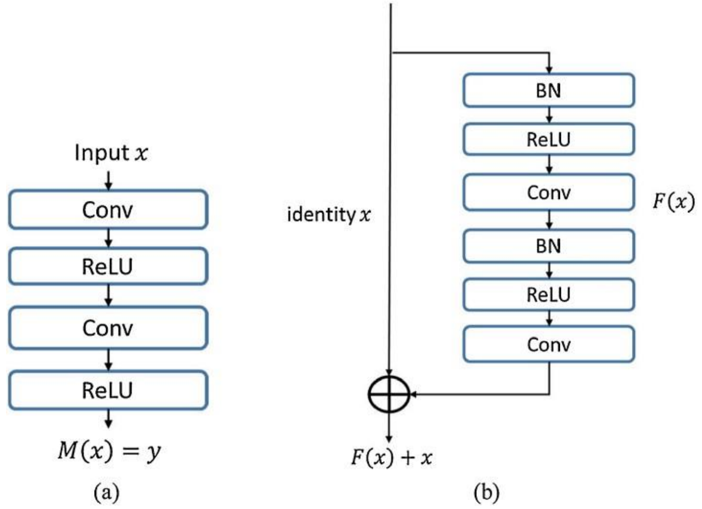
UNet

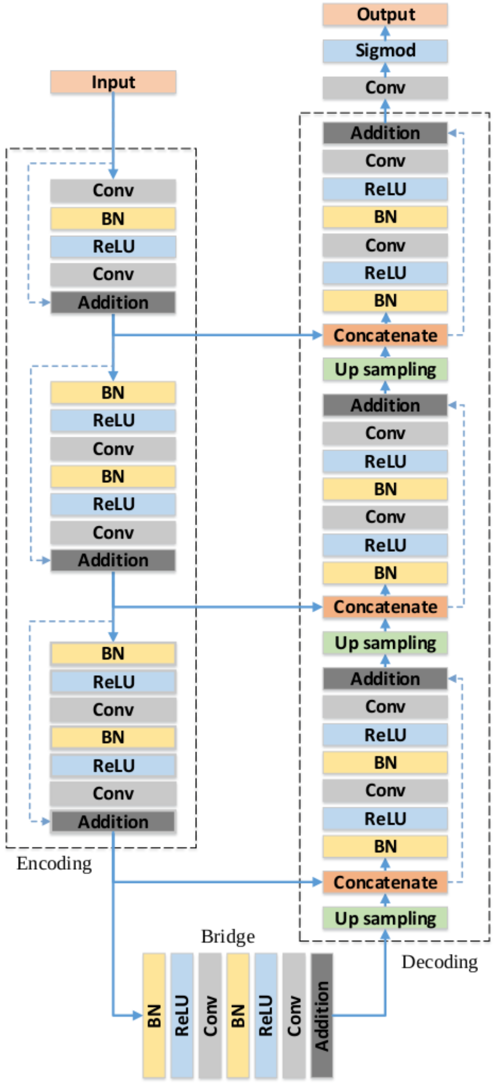


https://arxiv.org/abs/1505.04597

ResUNet

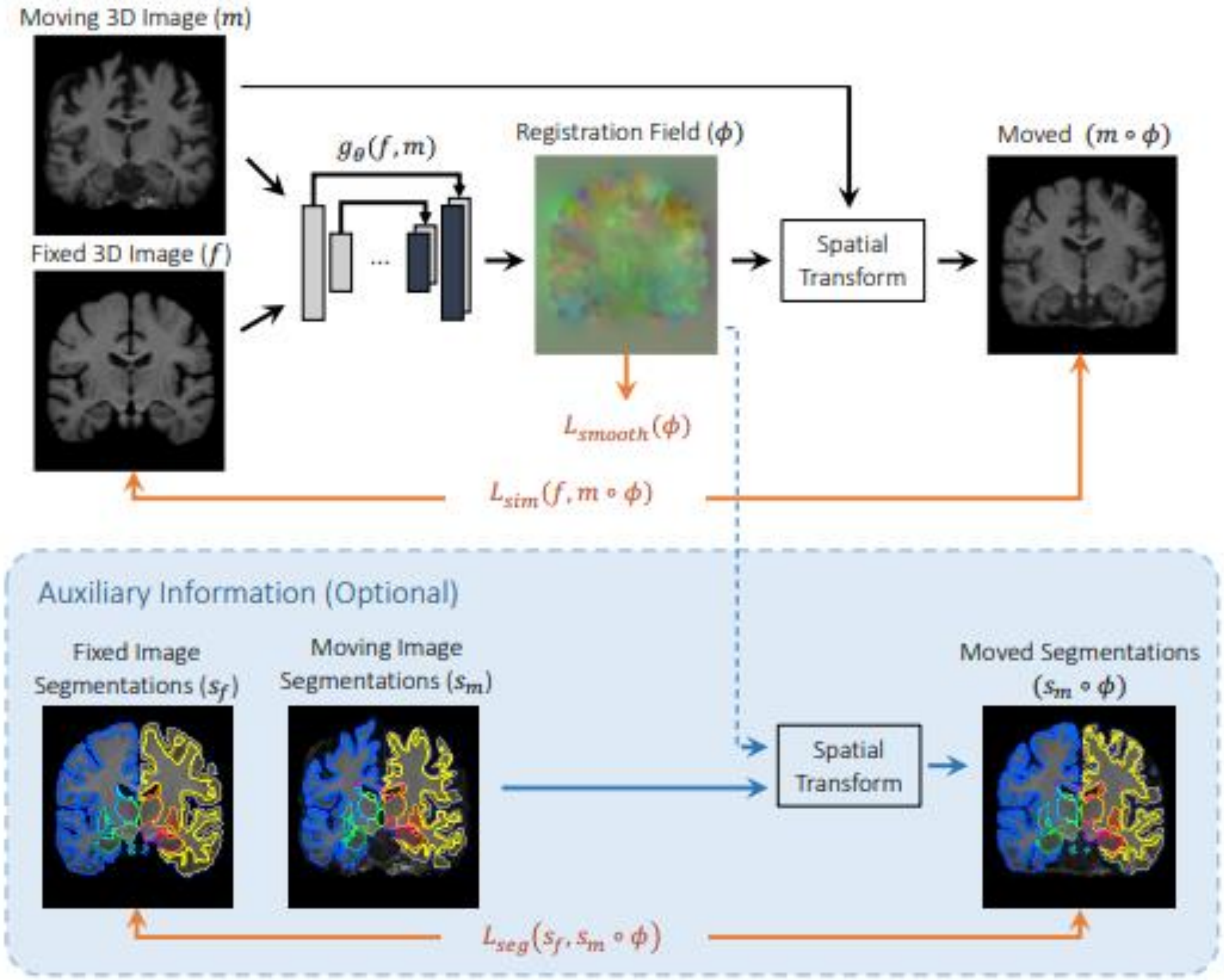


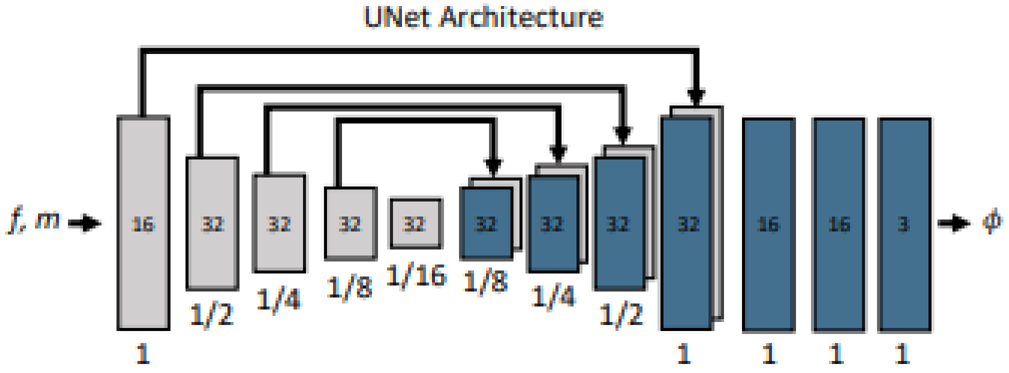




https://www.researchgate.net/figure/Architecture-of-the-residual-U-net-The-number-of-convolution-kernels-channels-is_fig1_330776159
https://medium.com/codex/architectures-for-medical-image-segmentation-part-3-residual-unet-ac5a4ca4212d
https://github.com/nikhilroxtomar/Deep-Residual-Unet/blob/master/Deep%20Residual%20UNet.ipynb

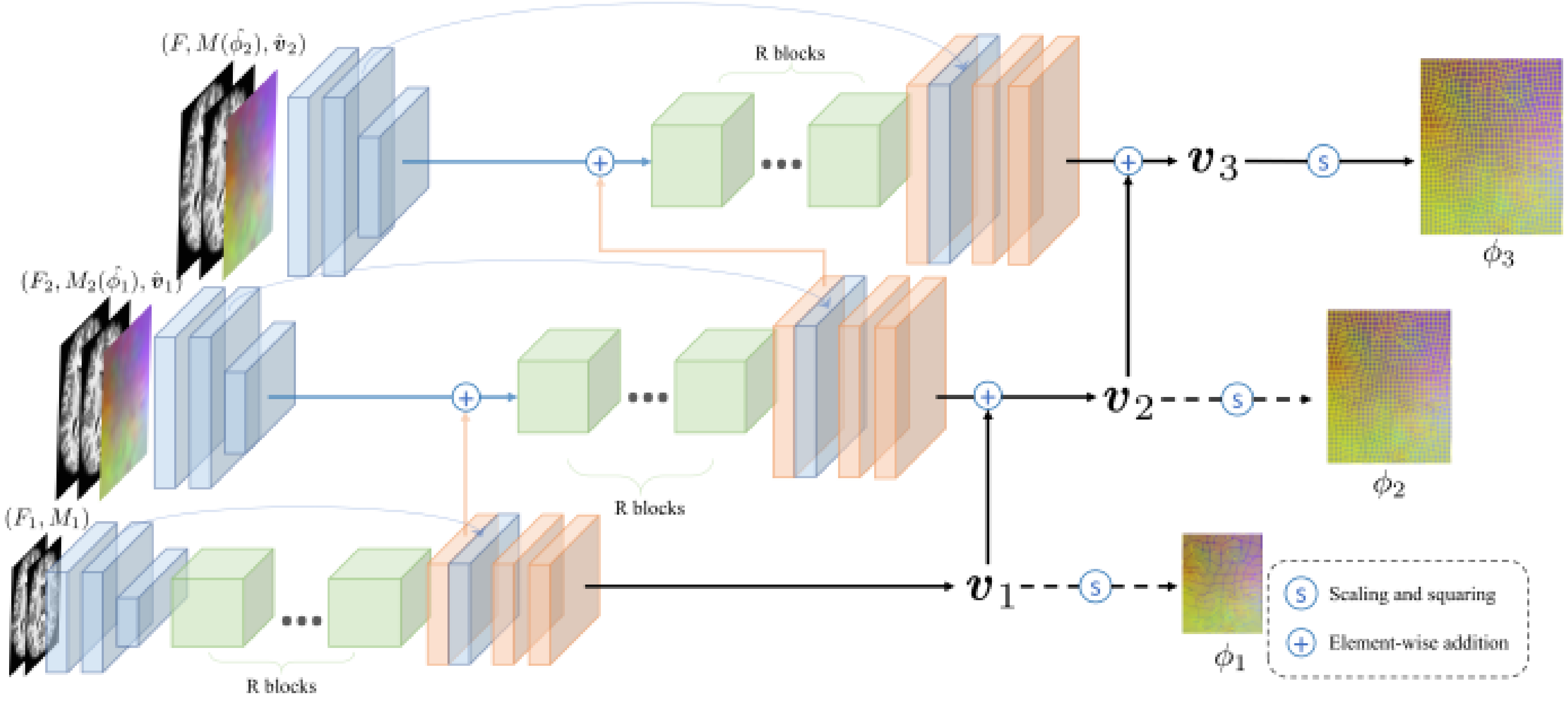
VoxelMorph





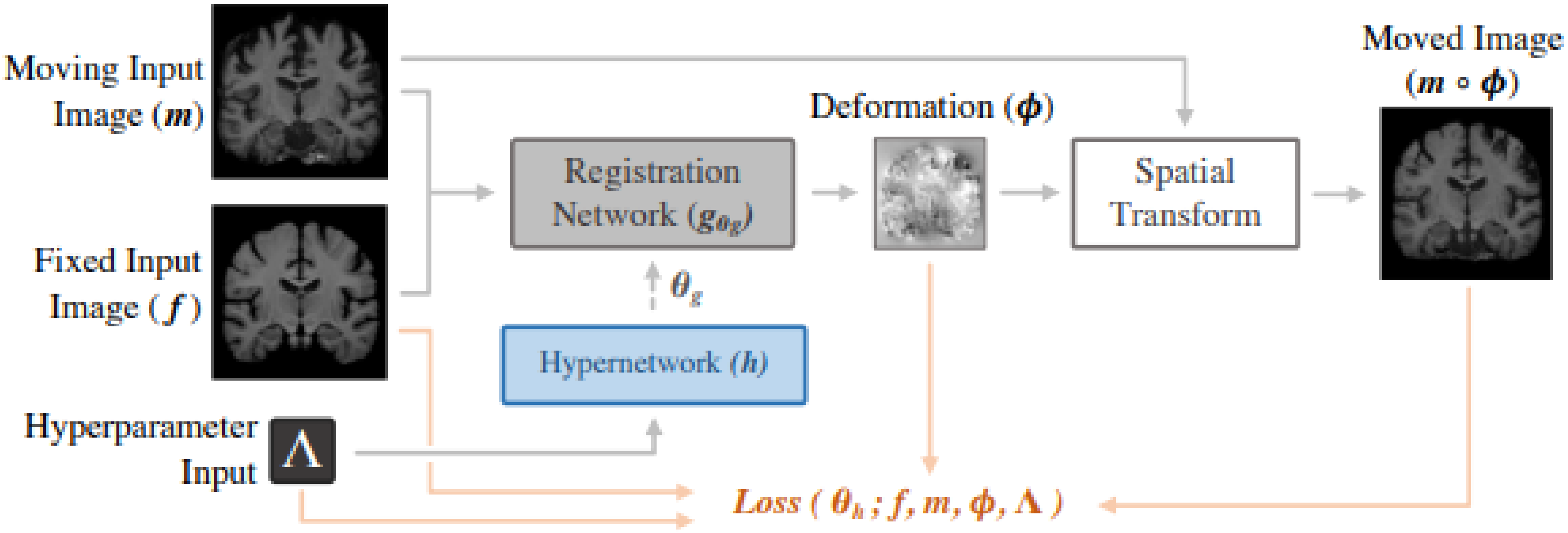
https://arxiv.org/pdf/1809.05231.pdf

LapIRN



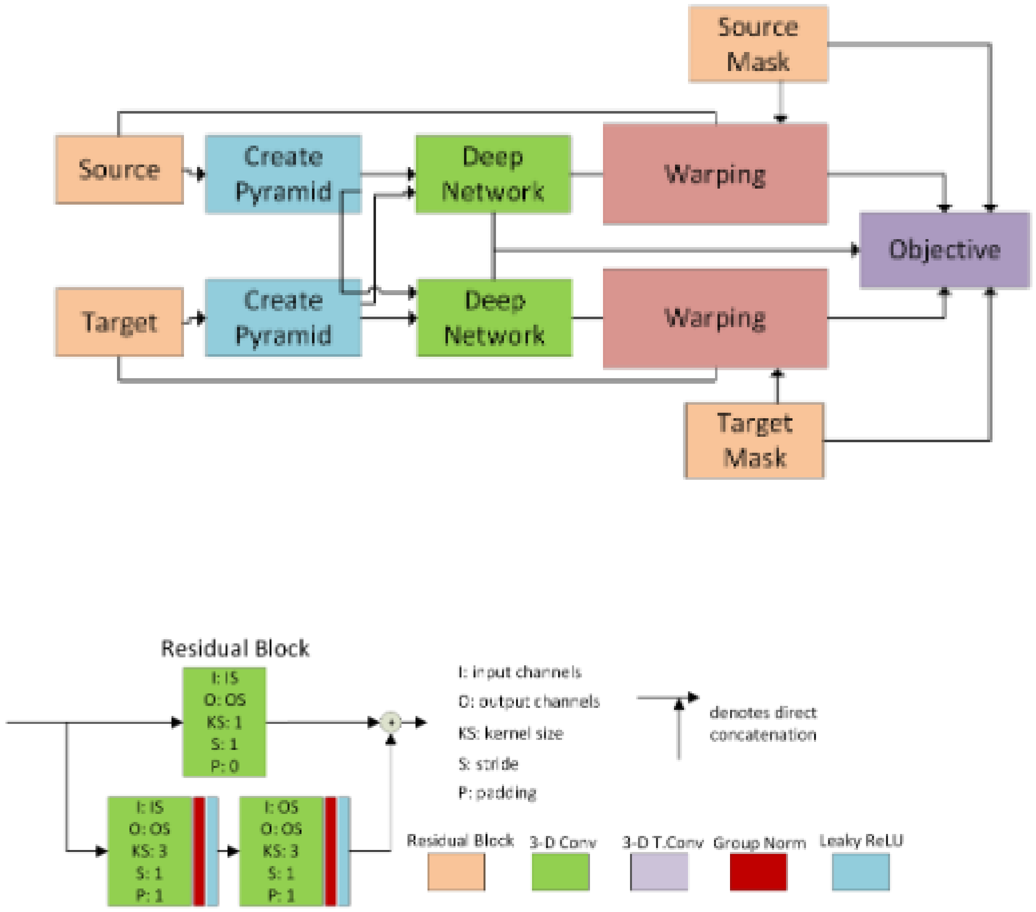
https://arxiv.org/pdf/2006.16148.pdf

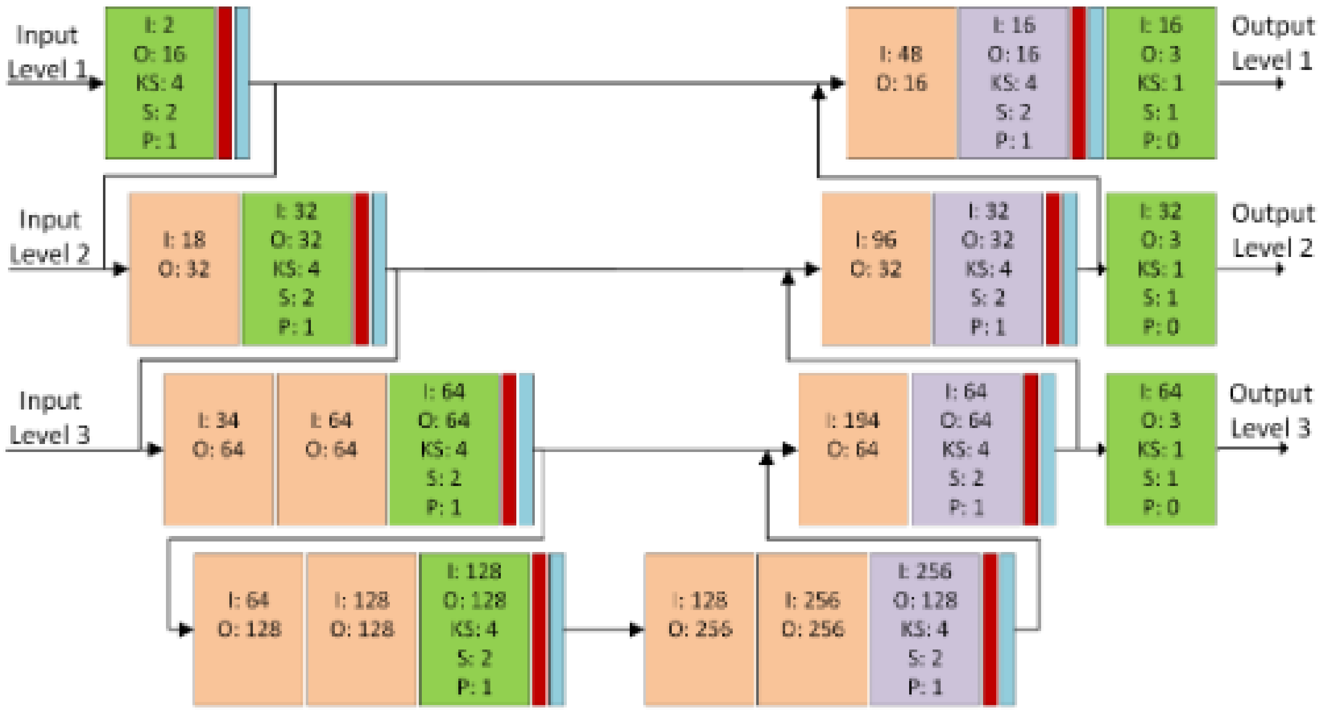
HyperMorph



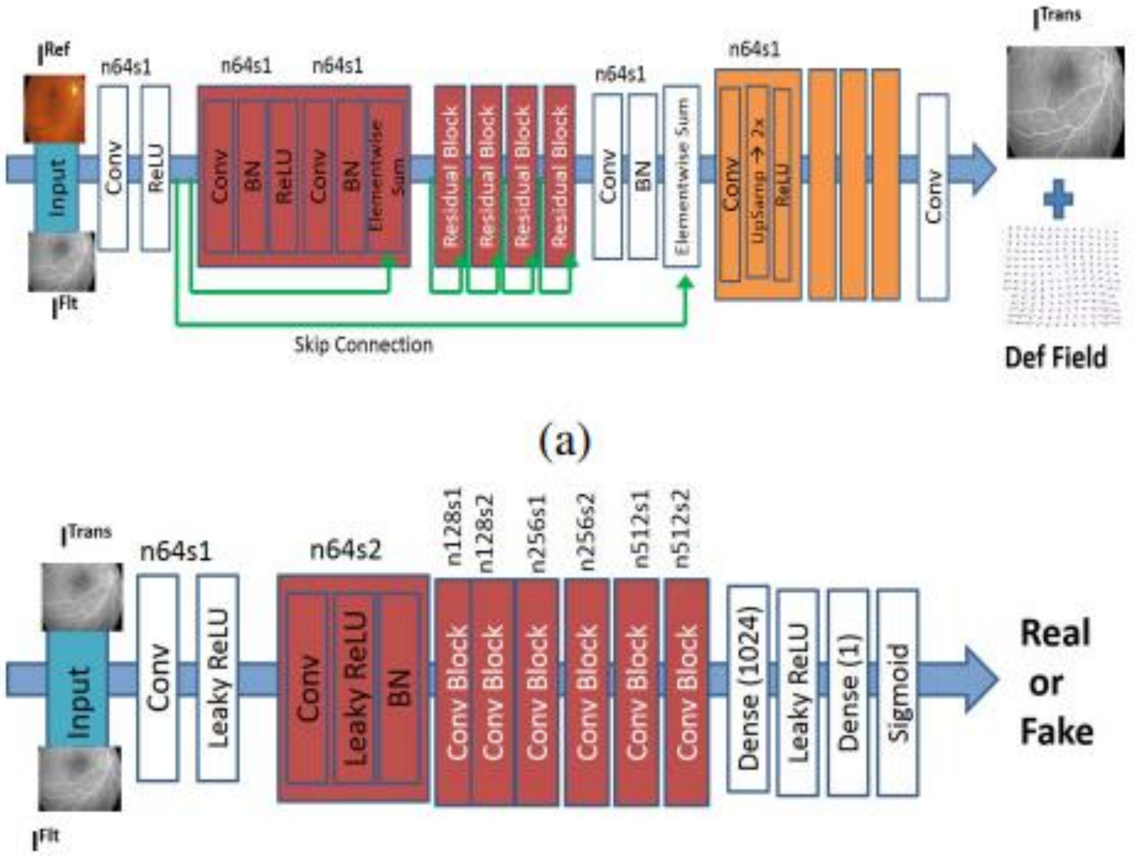
https://arxiv.org/pdf/2101.01035.pdf

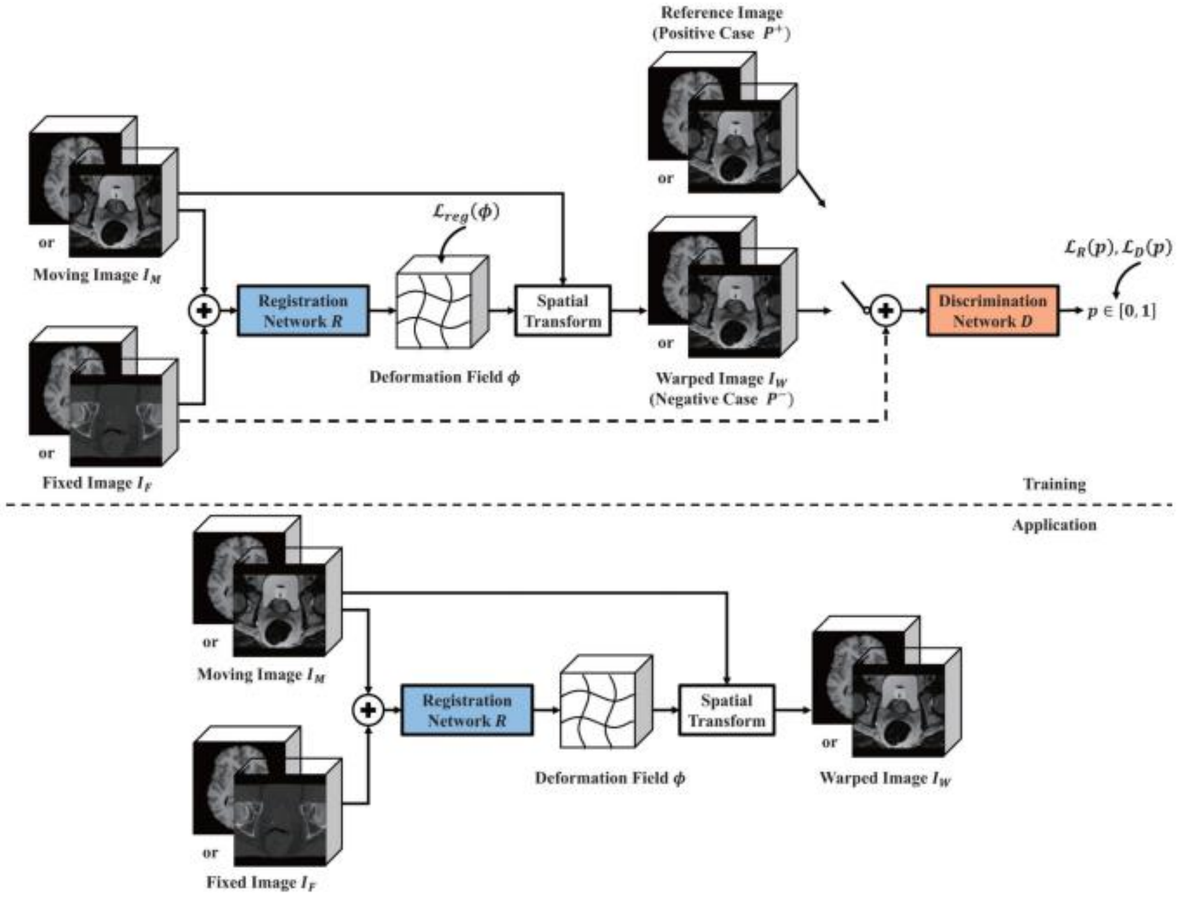
Symmetric, Multi-level, Resiudal UNet





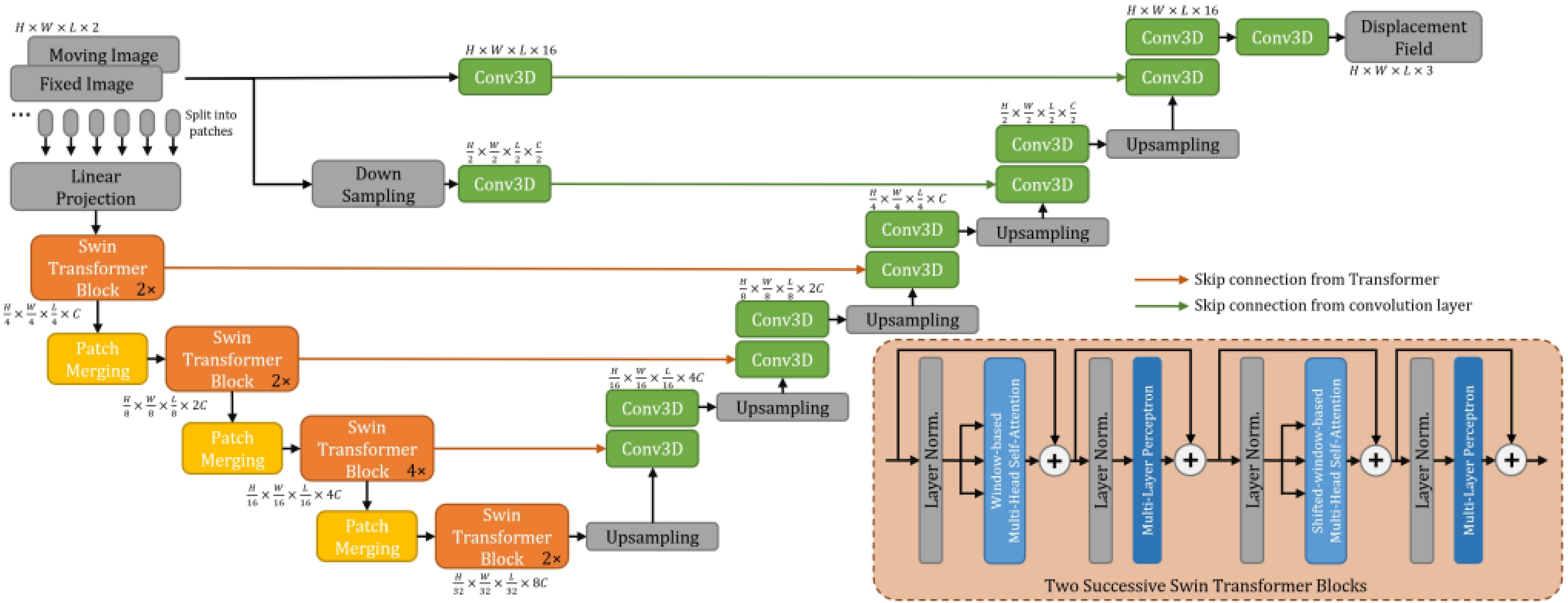
GAN-based Registration





https://arxiv.org/pdf/1805.02369.pdf
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7455790/pdf/nihms-1545621.pdf

TransMorph



https://arxiv.org/pdf/2111.10480.pdf

… and hundreds more

Evaluation of Medical Image Registration

AGH

RESEARCH UNIVERSITY
EXCELLENCE INITIATIVE

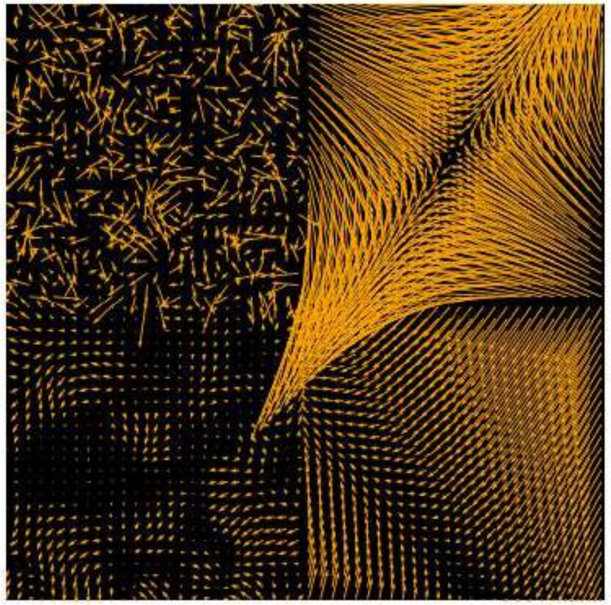
RMSE, CURL, Jacobian, Inverse Consistency

$$RMSE(y_1(\boldsymbol{x}), y_2(\boldsymbol{x})) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{D}(y_1(\boldsymbol{x}_{ij}) - y_2(\boldsymbol{x}_{ij}))^2},$$

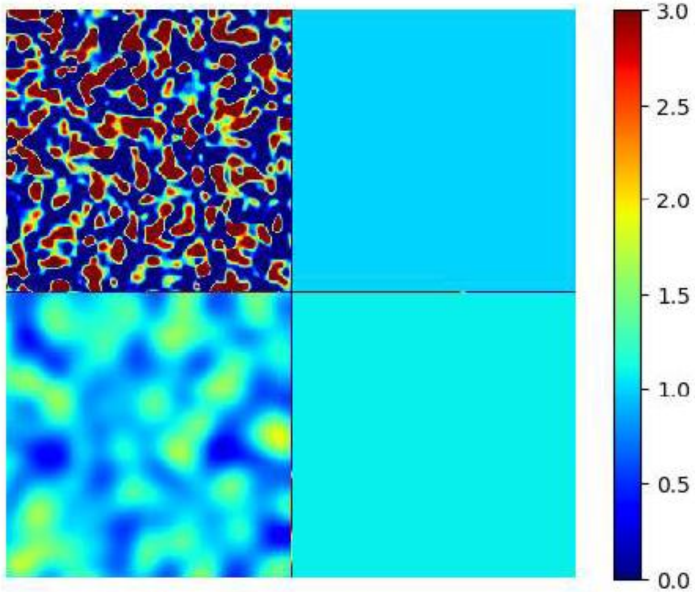
$$J(y(\boldsymbol{x})) = \begin{vmatrix} \frac{\delta y_1}{\delta x_1} & \frac{\delta y_1}{\delta x_2} & \frac{\delta y_1}{\delta x_3} \\ \frac{\delta y_2}{\delta x_1} & \frac{\delta y_2}{\delta x_2} & \frac{\delta y_2}{\delta x_3} \\ \frac{\delta y_3}{\delta x_1} & \frac{\delta y_3}{\delta x_2} & \frac{\delta y_3}{\delta x_3} \end{vmatrix},$$

$$CURL(y(\boldsymbol{x}_i)) = \nabla \times y(\boldsymbol{x}_i) = \begin{vmatrix} \hat{x_1} & \hat{x_2} & \hat{x_3} \\ \frac{\delta}{\delta x_1} & \frac{\delta}{\delta x_2} & \frac{\delta}{\delta x_3} \\ y(\boldsymbol{x}_{i1}) & y(\boldsymbol{x}_{i2}) & y(\boldsymbol{x}_{i3}) \end{vmatrix},$$

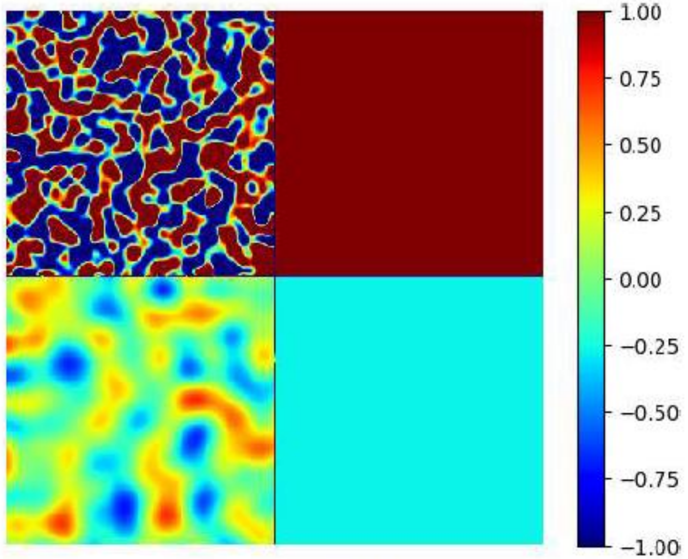
$$IC(y(\boldsymbol{x}), y^{-1}(\boldsymbol{x})) = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\sum_{j=1}^{D}((y \circ y^{-1})(\boldsymbol{x}_{ij}) - \mathrm{Id}(\boldsymbol{x}_{ij}))^2},$$



(a) Displacement Field



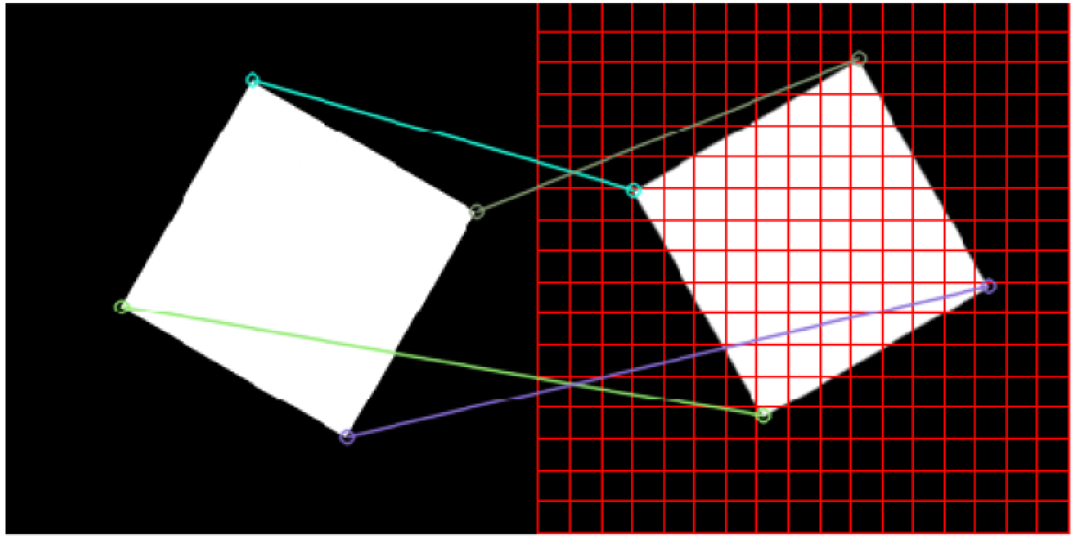
(b) Jacobian Determinant



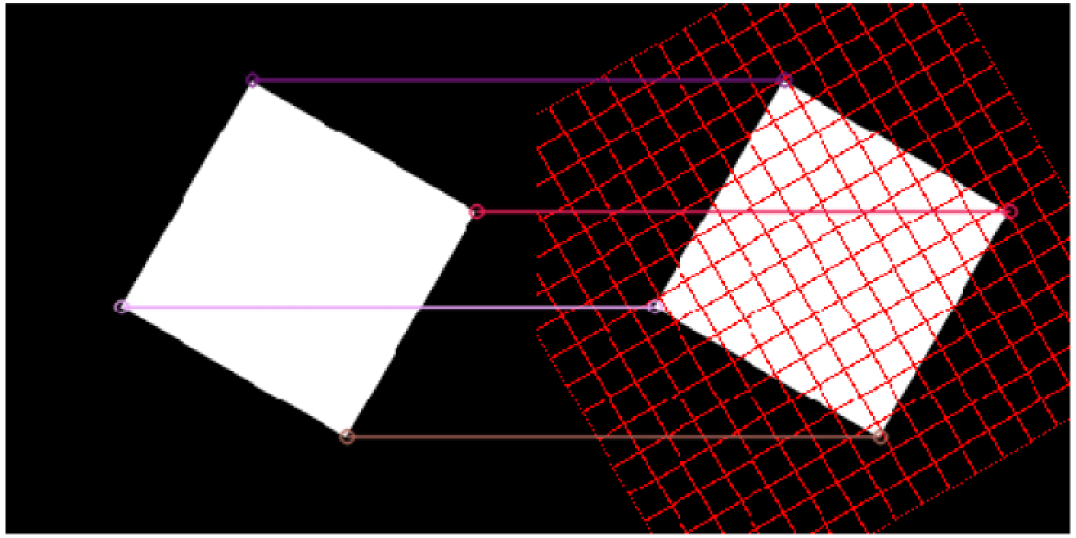
(c) Curl

Target Registration Error

$$TRE(\boldsymbol{q}, \boldsymbol{p}) = \sqrt{\sum_{i=1}^{D} (q_i - p_i)^2},$$

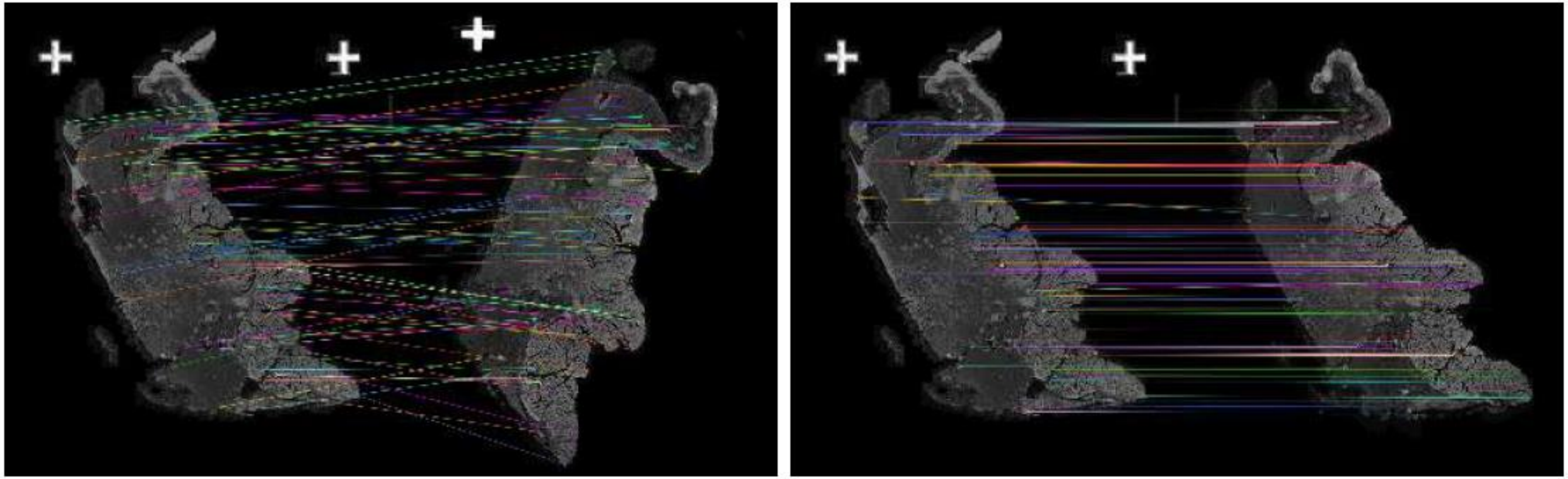


(a) Before registration



(b) After registration

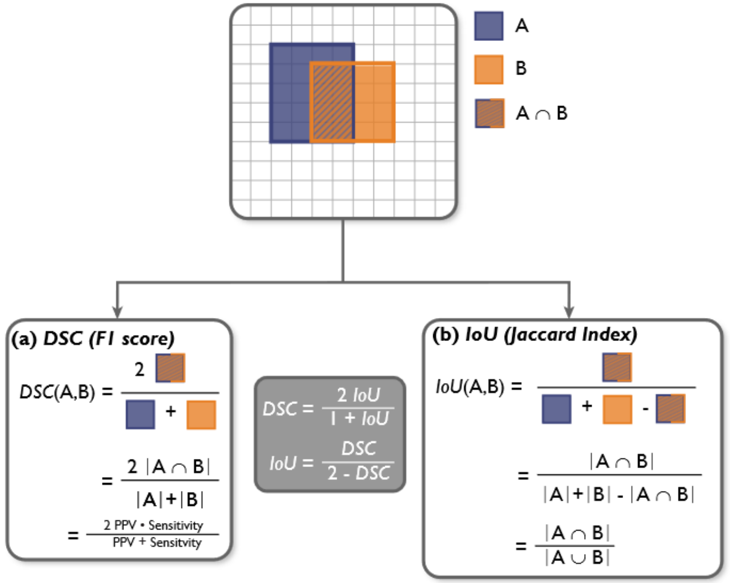
Target Registration Error

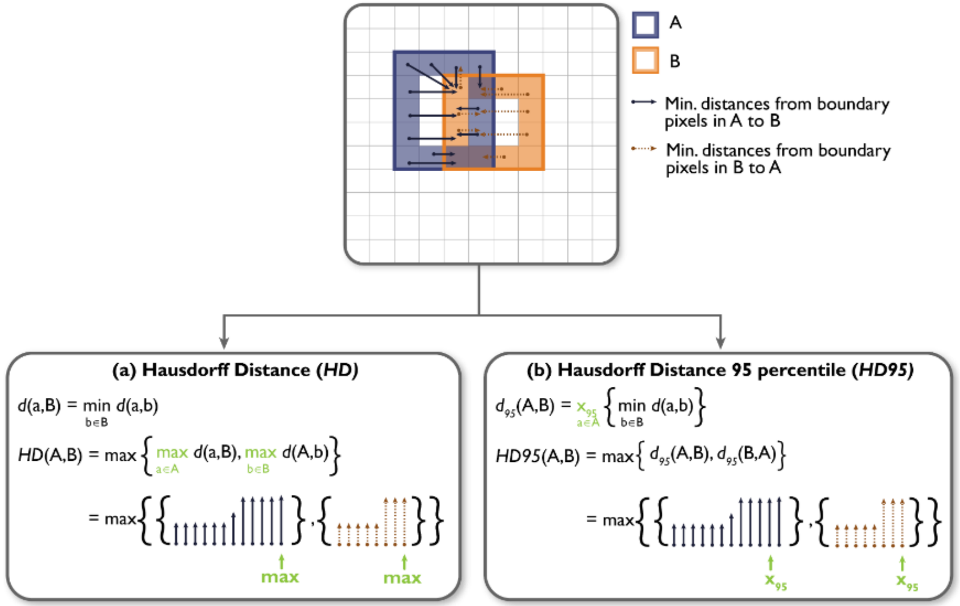


(a) Before the registration (b) After the registration

Dice / Hausdorff

$$DICE(M_b(\boldsymbol{x}), F_b(\boldsymbol{x})) = 2\frac{|M_b(\boldsymbol{x}) \cap F_b(\boldsymbol{x})|}{|M_b(\boldsymbol{x})| + |F_b(\boldsymbol{x})|},$$



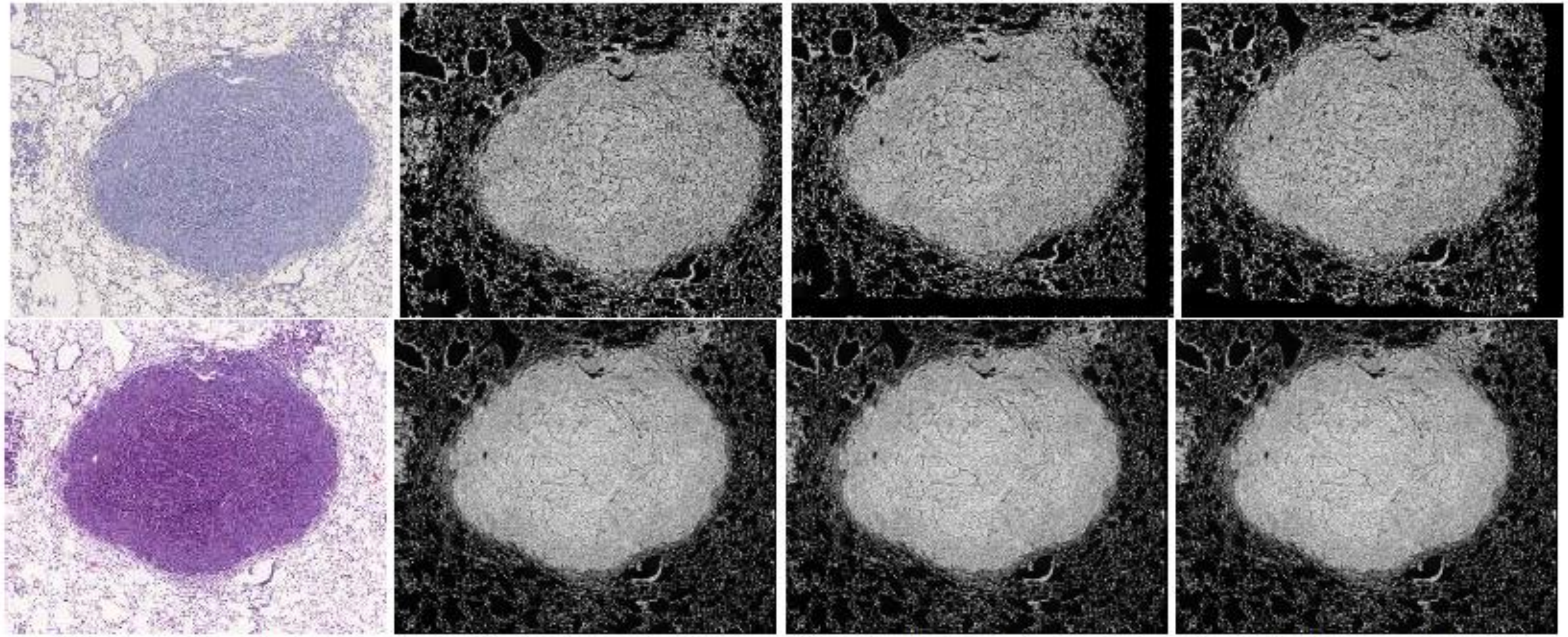


$$HD(M_b(\boldsymbol{x}), F_b(\boldsymbol{x})) = max(hd(M_b(\boldsymbol{x}), F_b(\boldsymbol{x})), hd(F_b(\boldsymbol{x}), M_b(\boldsymbol{x}))),$$

$$hd(M_b(\boldsymbol{x}), F_b(\boldsymbol{x})) = \max_{x_m \in M_b(\boldsymbol{x})} \left(\min_{x_f \in F_b(\boldsymbol{x})} (d(x_m, x_f)) \right),$$

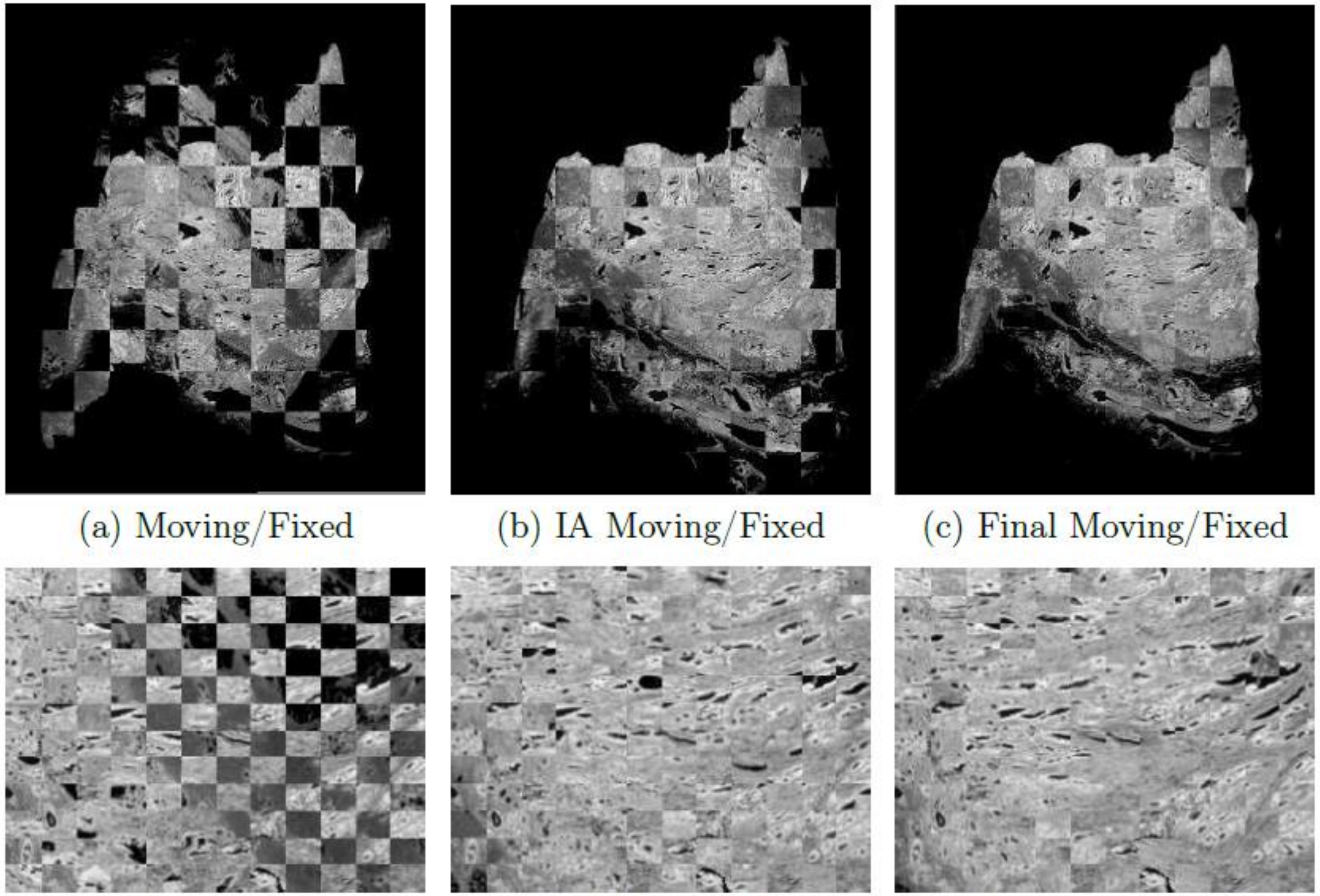
https://arxiv.org/abs/2104.05642

Visual Assessment



a) Original b) Preprocessed c) Affinely Registered d) Final result

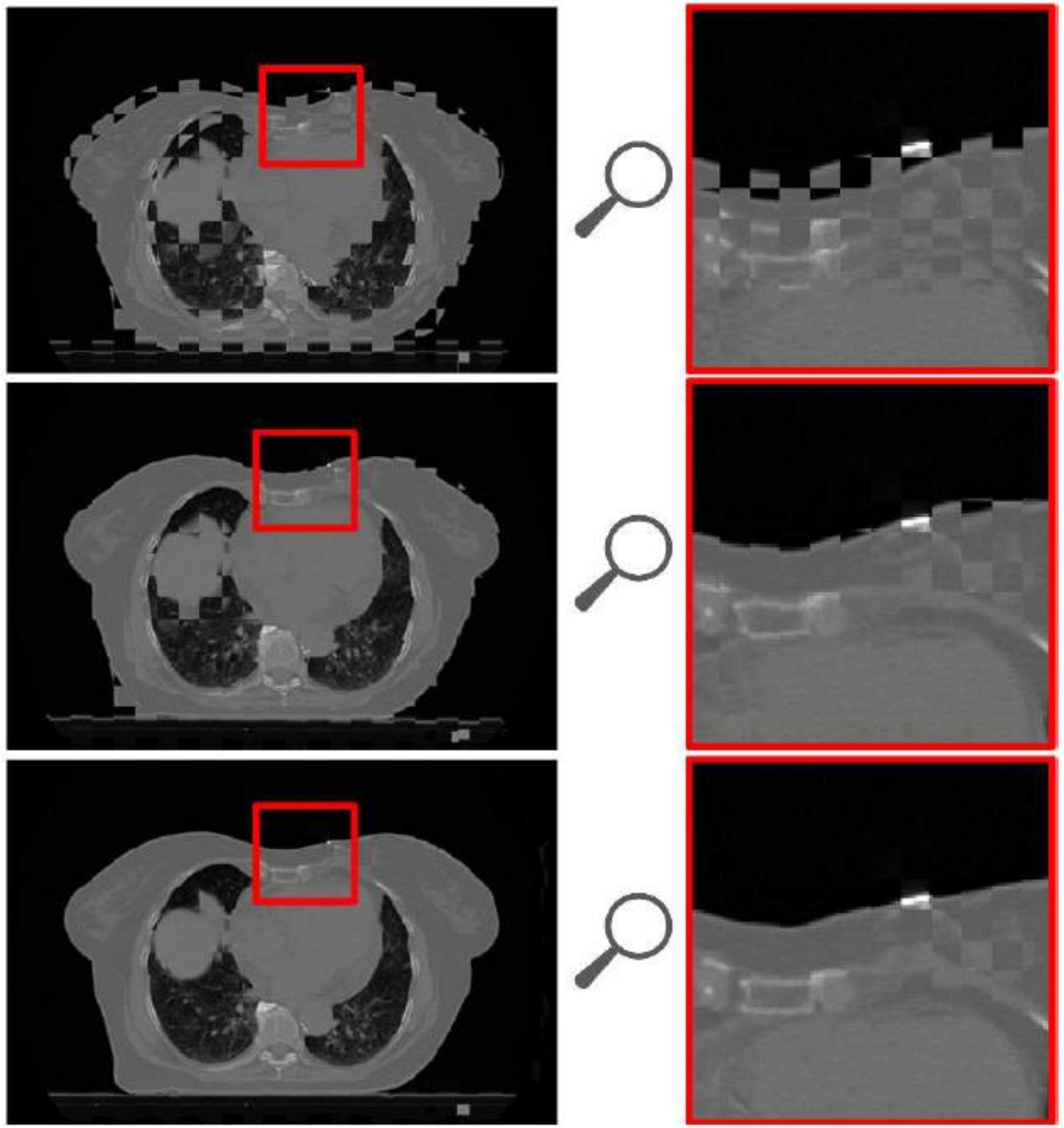
Visual Assessment



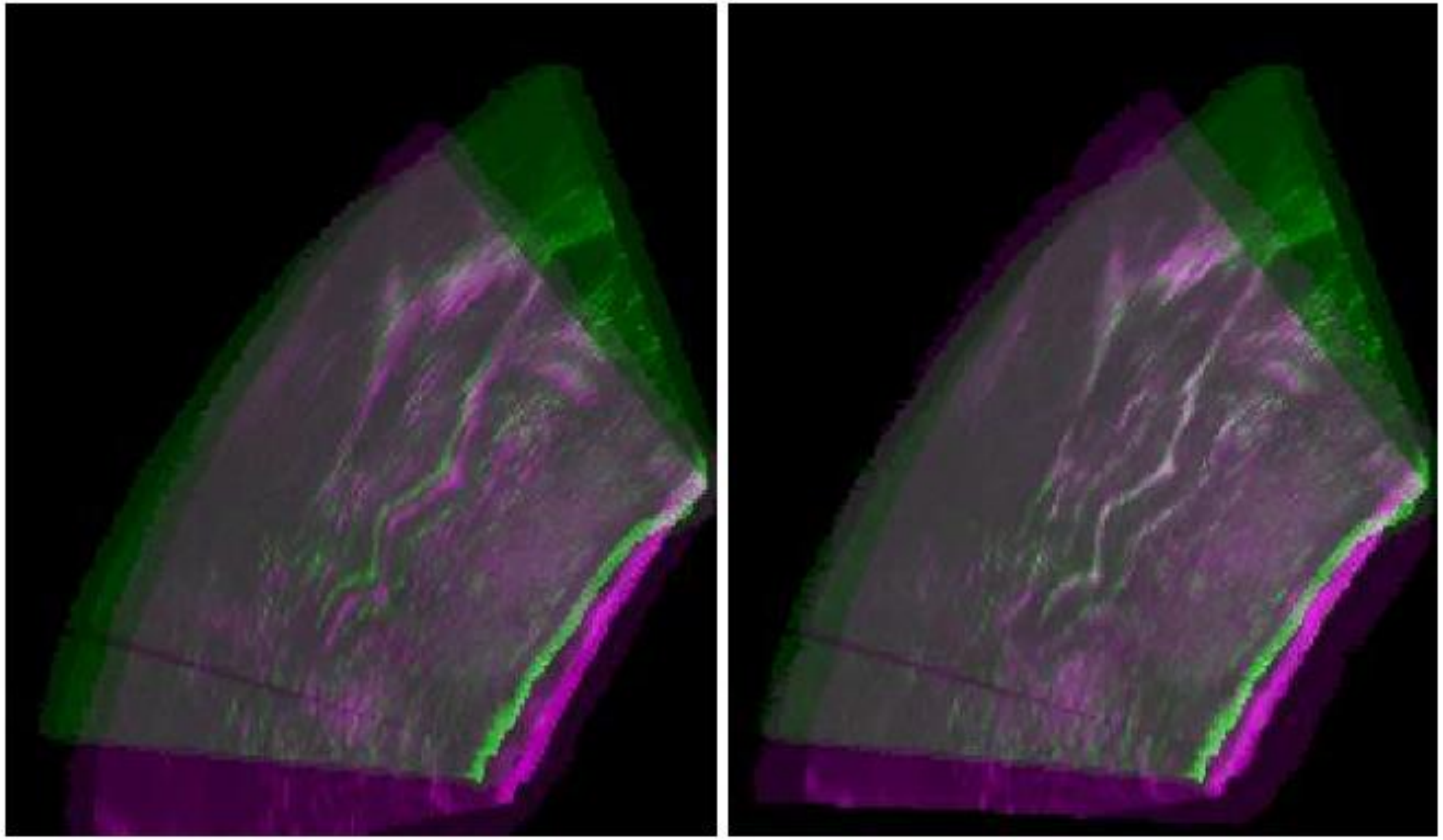
(a) Moving/Fixed (b) IA Moving/Fixed (c) Final Moving/Fixed

(d) Moving/Fixed (e) NR Moving/Fixed (f) NR Final Moving/Fixed

Visual Assessment



Visual Assessment



Questions?

Practical

- How to transform images?
- Lab 7 - introduction

AGH

RESEARCH UNIVERSITY
EXCELLENCE INITIATIVE

15 min break