Technische Universiteit
Eindhoven
University of Technology

# Deep learning for (deformable) medical image registration

Maureen van Eijnatten

# Lecture schedule

| Week | Date | Lecturer | Topics |
|---|---|---|---|
| 1 | 1 Sept. | Maureen | Course introduction; Software demo; Image registration (1) |
| | 3 Sept. | Maureen | Image registration (2); Geometrical transformations |
| 2 | 8 Sept. | Maureen | Point-based registration |
| | 10 Sept. | Maureen | Intersity-based registration; Evaluation metrics |
| 3 | 15 Sept. | *Catch-up day (no lecture)* | |
| | 17 Sept. | Cornel Zachiu (UMCU) | Guest lecture 1: Image analysis for adaptive radiotherapy |
| 4 | 22 Sept. | Mitko | Introduction to CAD; k-NN; Decision trees |
| | 24 Sept. | Mitko | Generalization and overfitting |
| 5 | 29 Sept. | Mitko | Logistic regression; Neural networks |
| | 1 Oct. | Friso | Convolutional neural networks |
| 6 | 6 Oct. | Friso | Deep learning frameworks and applications |
| | 8 Oct. | Friso | Unsupervised machine learning |
| **7** | **13 Oct.** | **Maureen** | **Deep learning for deformable image registration** |
| | 15 Oct. | Geert-Jan Rutten (ETZ) | Guest lecture 2: Image analysis in neurosurgery applications |
| 8 | 20 Oct | *Self-study (no lecture)* | Active shape models |
| | 22 Oct | *Self-study (no lecture)* | Active shape models |

**Outline for today:**

- Quick recap of medical image registration
  - Rigid, affine, and non-linear registration models
  - Introduction non-parametric / deformable image registration

- Deep learning
  - Deep iterative registration
  - Supervised methods
  - Unsupervised (optimization-based) methods

- Feedback on project 1

- Question hour and example exam questions

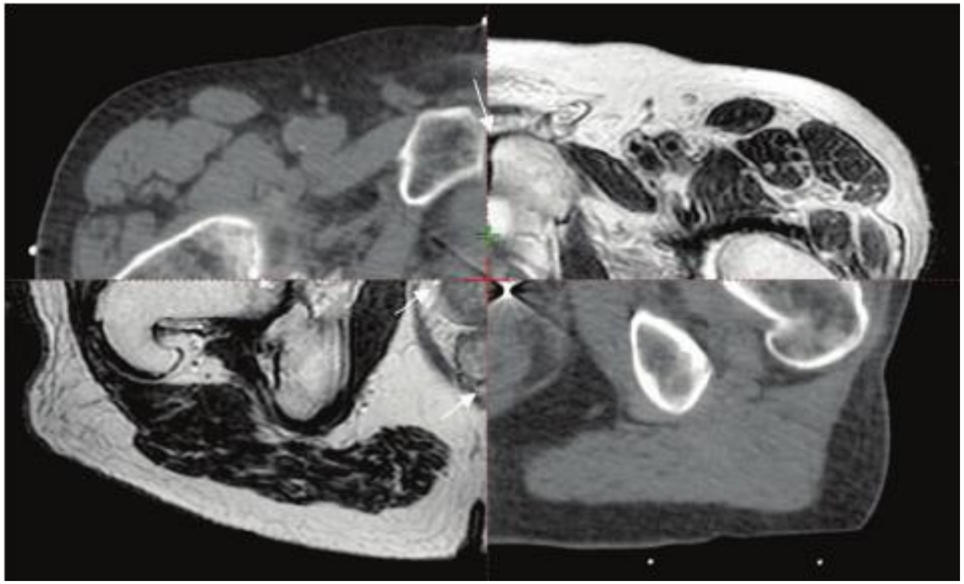
## Medical image registration

Why important?

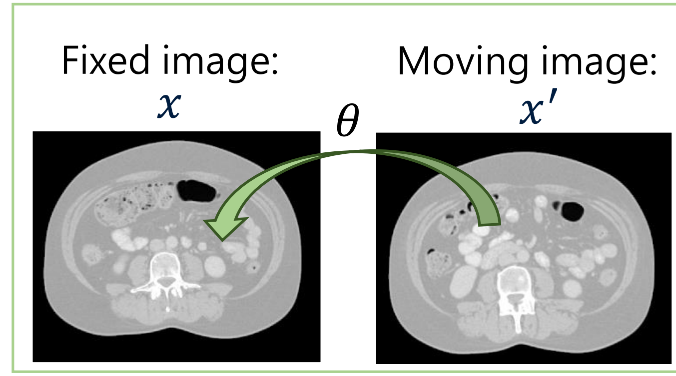
- Multi-modal registration (e.g. CT on MRI)
- Inter-subject (e.g. atlas registration)
- Longitudinal (e.g. treatment evaluation)

*Challenges*

- Large 3D volumes
- Accuracy vs. efficiency
- Intensity inhomogeneities and discontinuities
- Outlier rejection



# Medical image registration: general "recipe"




**Transformation model:**
(e.g., rigid, affine, deformable)

$$T(x', \theta)$$

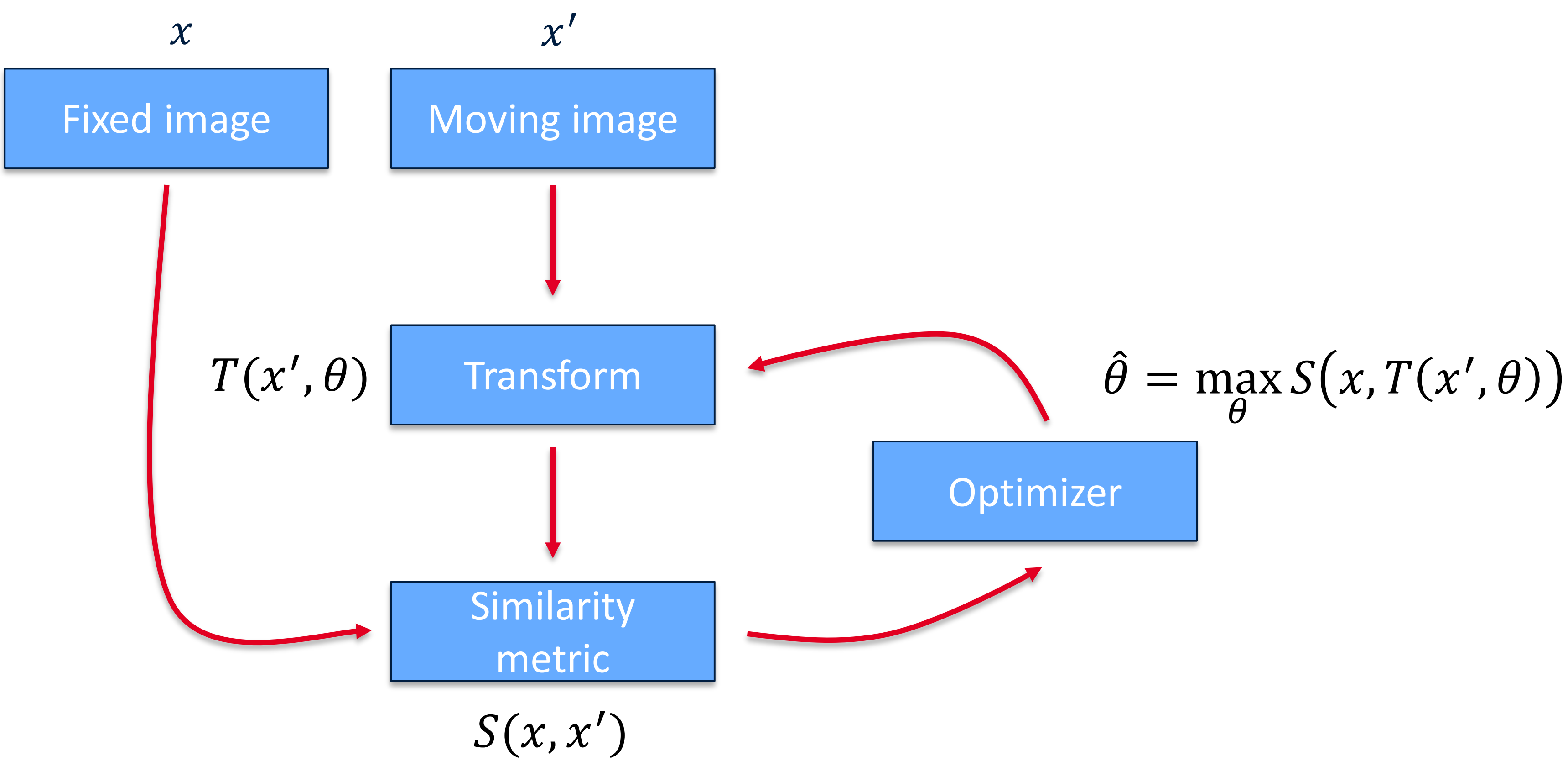

**Similarity measure:**
(e.g. SSD, CC, MI, MSE)

$$S(x, x')$$

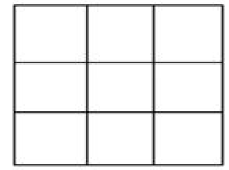
**Optimization:**

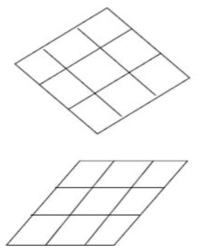
$$\hat{\theta} = \max_{\theta} S\big(x, T(x', \theta)\big)$$

## Finding the optimal transformation

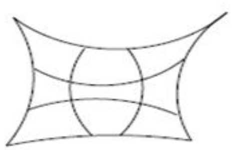


# Transformation models





**Rigid & Affine**
- Point or feature-based (e.g. landmarks, fiducials)
- Intensity-based
- Gradient/edge-based



**Non-linear**
- Linear or higher order polynomials
- Spline-based

## Rigid transformations

$$\mathbf{x}' = \mathbf{R}\mathbf{x} + \mathbf{t}$$

- Translation

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

- Rotation

$$\mathbf{R} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$$

## Affine transformations

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{t}$$

Translation, rotations and:

- Scaling

$$\mathbf{S} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$$

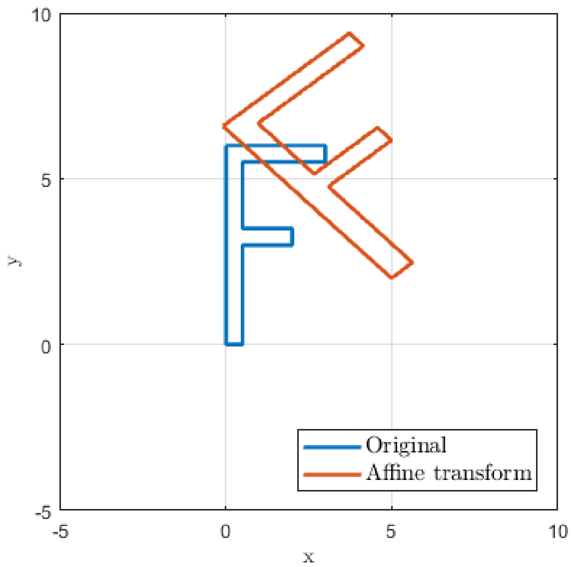
- Shearing

$$\mathbf{H} = \begin{bmatrix} 1 & h_x \\ h_y & 1 \end{bmatrix}$$

## Affine transformation (no restriction on the transformation parameters):

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{t}$$



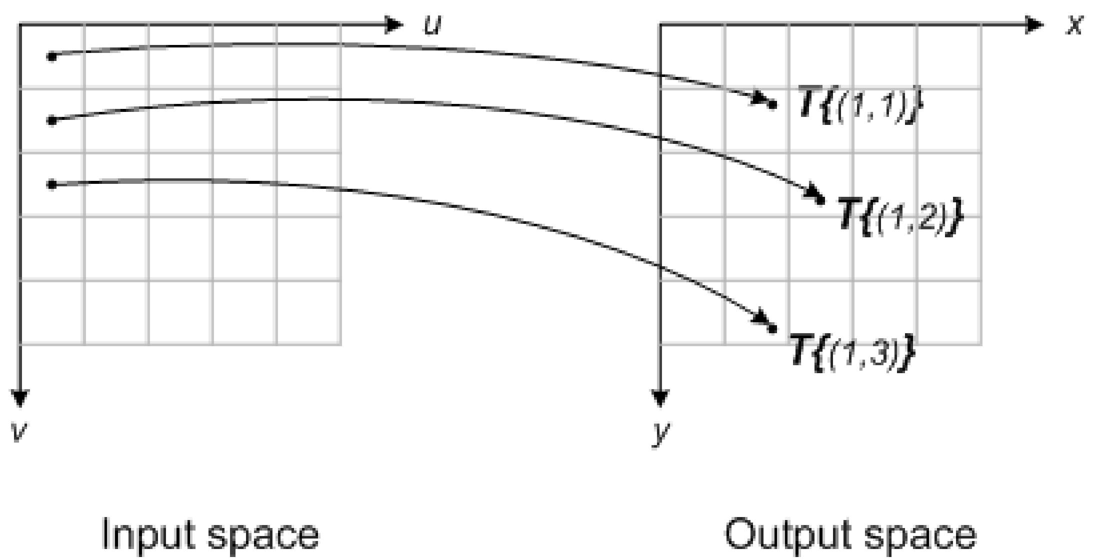

Can be considered as a composition of any combination of rotations, scalings, shearings, reflections + translations.

How many parameters (in 2D)?

And in 3D?

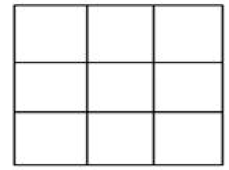
3D: $$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c & t_x \\ d & e & f & t_y \\ g & h & i & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

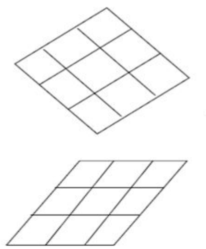
Transforming an image means transforming the spatial coordinates of the pixels.



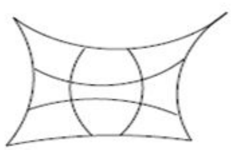

So far, we have parameterized our transformation model by constraining it to a rigid, affine or non-linear transformation. Is it also possible to define a vector for every pixel?

# Transformation models

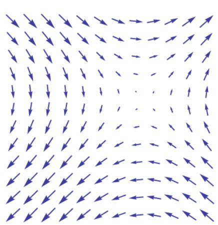




## Rigid & Affine
- Point or feature-based (e.g. landmarks, fiducials)
- Intensity-based
- Gradient/edge-based

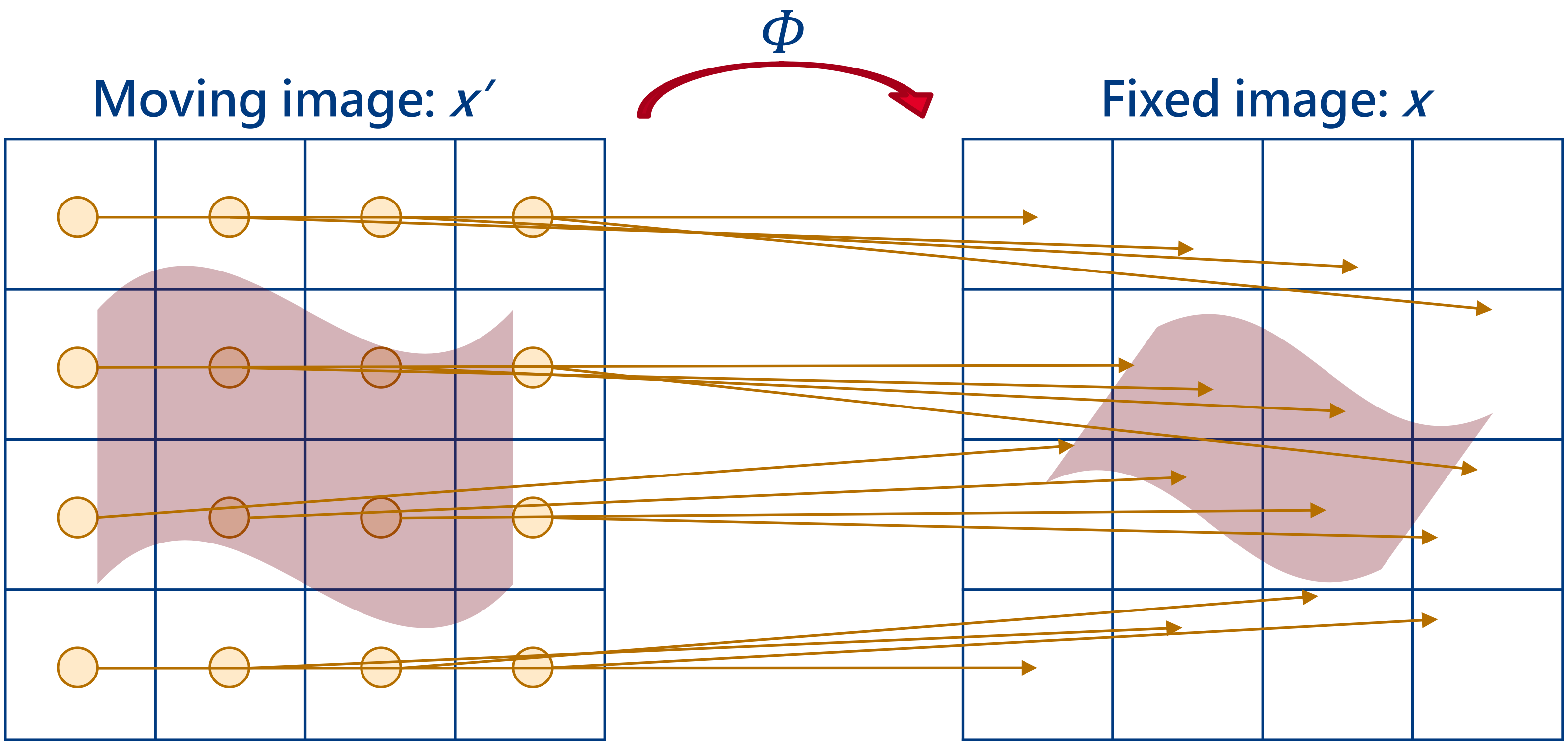


## Non-linear
- Linear or higher order polynomials
- Spline-based



## Non-parametric / deformable
- Allowing each image element to be displaced arbitrarily

**Displacement (vector) field (DVF)** = Dense set of vectors representing the displacement in a given spatial domain



# Formulation of deformable image registration

Images (fixed image $x$, moving image $x'$): $x : \Omega \to \mathbb{R}$

Image domain: $\Omega \to \mathbb{R}^d, d = 2,3$

Transformed moving image:

Regularization

$$u' = \underset{u \in H}{\arg\min}\; E_D\left(x, \boxed{x'(\varphi)}\right) + \lambda\, E_R(u)$$

Difference measure

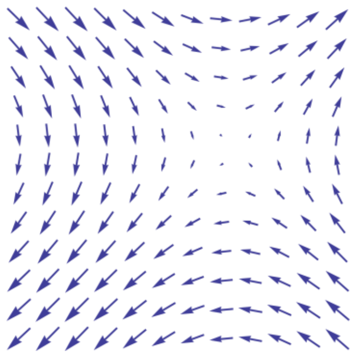
Deformation: $\varphi = \mathrm{Id} + u,\quad \varphi : \Omega \to \mathbb{R}^d$

or point-wise: $\varphi(x) = x + u(x)$

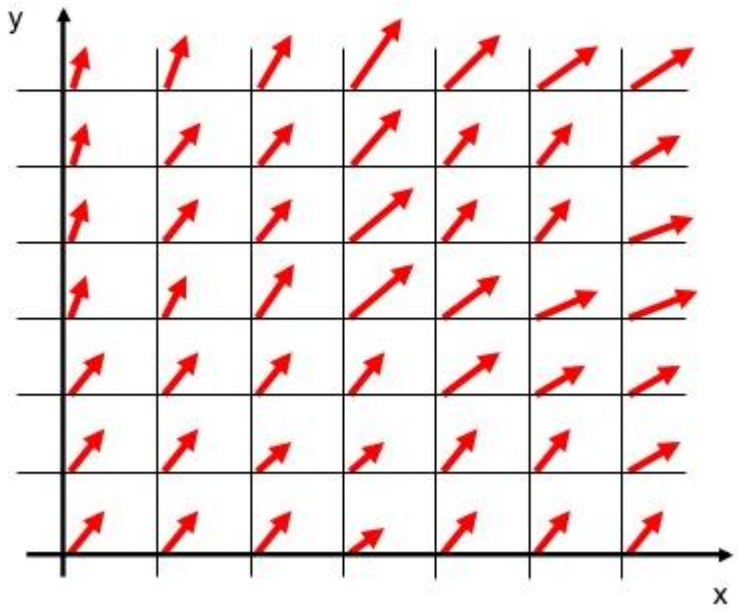
Displacement:

$u : \Omega \to \mathbb{R}^d$

e.g. $u = [u_x, u_y, u_z]$



# Deformable image registration: how many free parameters (DOF)?




Typical spatial resolution of a
3D medical image
(MRI: e.g. $1.5 \times 1.5 \times 4$ $mm^3$ voxels)


DOF = 3*Nx*Ny*Nz (!)

Deformable image registration is still a very active area of research, and many different deformable registration models exist:

- Free-form deformation model
- Optical flow
- Demons
- Fluid flow
- Diffeomorphisms
- …

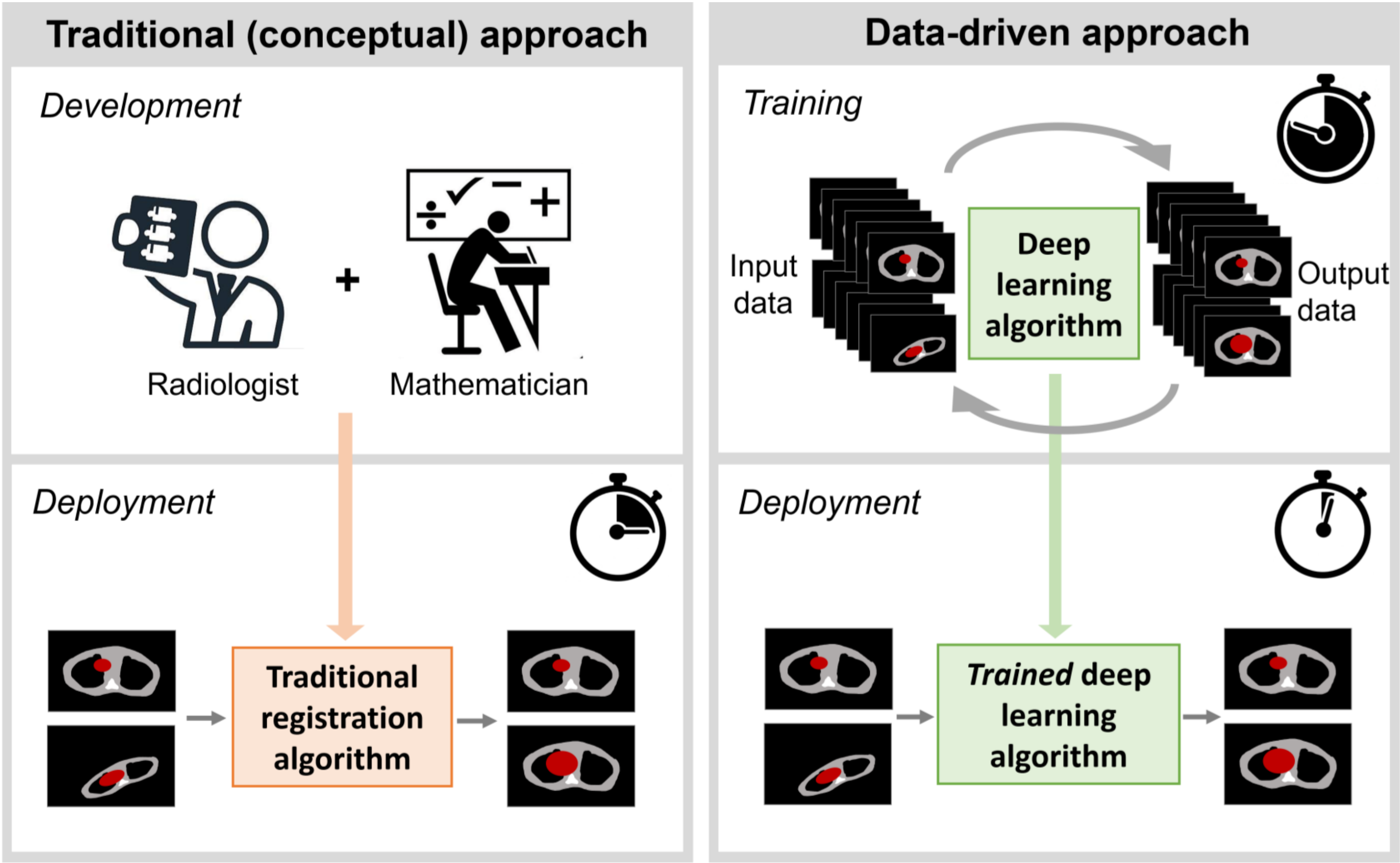
Note that the details of these models and their implementations are beyond the scope of this course.

## Learning goals of this lecture

The students should be able to:

- Explain the difference between parametric and non-parametric image registration models
- Formulate deformable image registration as an optimization problem
- Understand displacement vector fields
- Explain the difference between supervised and unsupervised learning
- Understand why deep learning is an interesting technique to solve medical image registration tasks
- Understand the different ways in which deep learning can be used to perform (deformable) image registration

# Why focus on deep learning for medical image registration?



# Machine learning – different training strategies

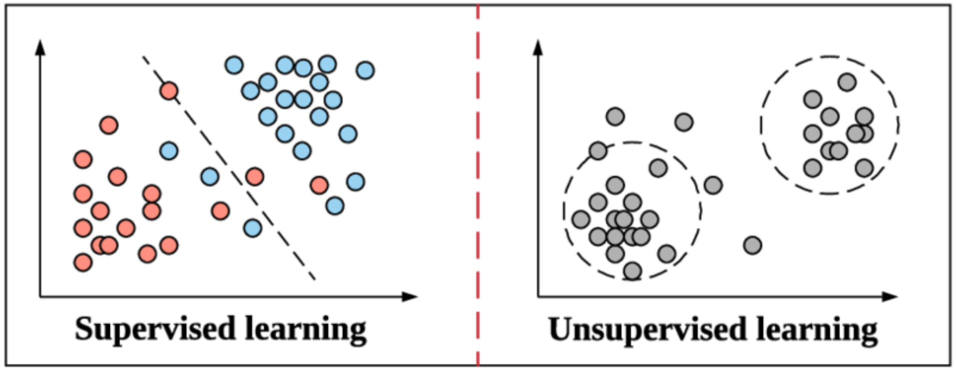
**Supervised learning** = develop predictive model based on input and output data (i.e., ground truth, "labels")

Examples: classification (e.g., skin lesion classification), regression (segmentation of vessel structures)

**Unsupervised learning** = group or interpret data based on input data alone

Example: clustering (e.g., k-means)

*NB: also semi-supervised learning and reinforcement learning, not part of this course.*



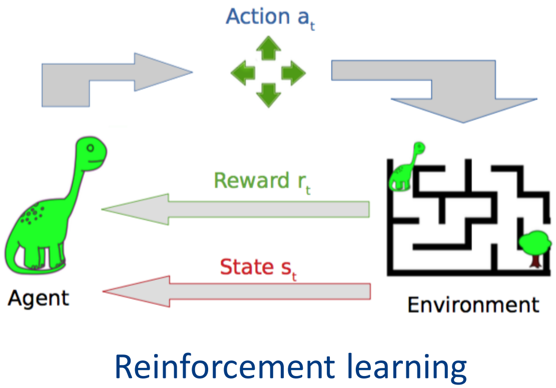






Image 1 source: Qian et al. (2019). Orchestrating the Development Lifecycle of Machine Learning-Based IoT Applications: A Taxonomy and Survey.
Image 2 source: https://twimlai.com/whats-hot-in-ai-deep-reinforcement-learning-edition/

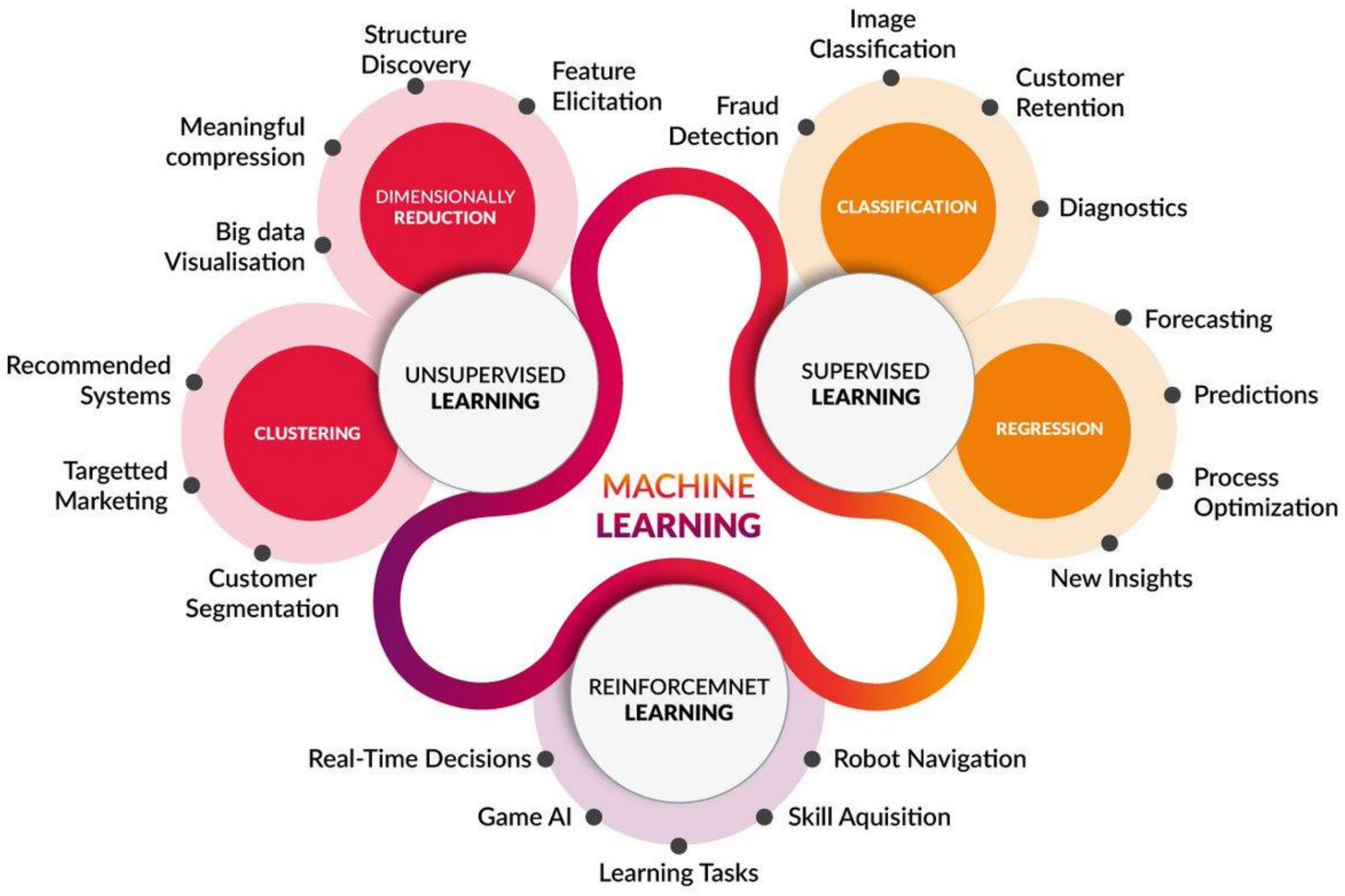




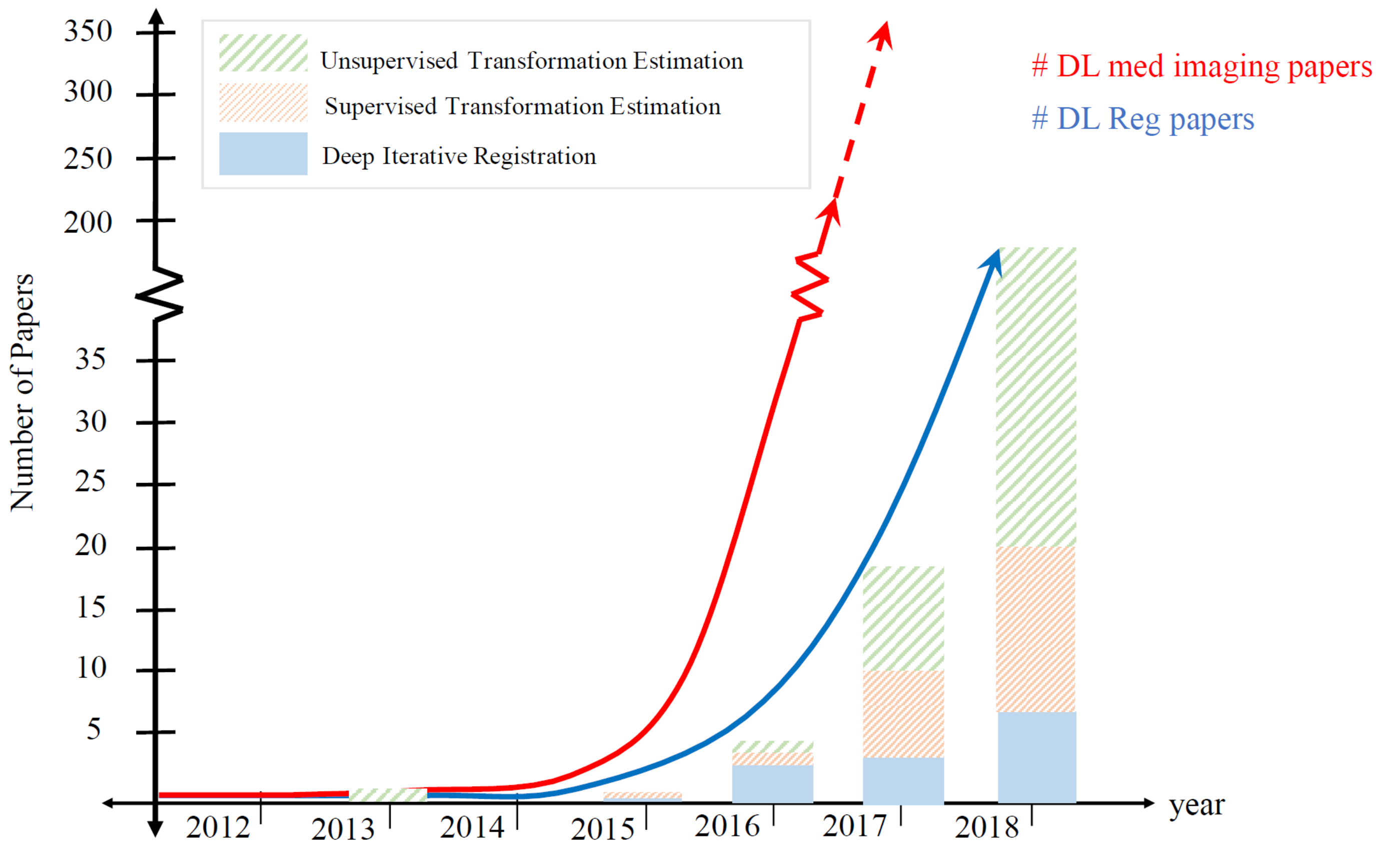

Image source: http://www.cognub.com/index.php/cognitive-platform/

Deep learning-based image registration methods for medical image registration can be subdivided as follows:

A. Deep iterative registration

B. Supervised transformation estimation

C. Unsupervised transformation estimation






Haskins et al. (2019) Deep Learning in Medical Image Registration: A Survey. arXiv:1903.02026v1

## A. Deep iterative registration

Use a traditional registration method and learn one component

### Automatically extract (learn) similarity features

- *Unimodal*: Wu et al. (2013) Unsupervised deep feature learning for deformable registration of mr brain images (MICCAI); Wu et al. (2016) Scalable high-performance image registration framework by unsupervised deep feature representations learning (IEEE Transactions on Biomedical Engineering); Eppenhof et al. (2018) Error estimation of deformable image registration of pulmonary ct scans using convolutional neural networks (Journal of Medical Imaging)
- *Multimodal*: Mostly *rigid registration*; learn a similarity metric to evaluate or register multimodal images using gradient descent. E.g. Cheng et al. (2016&2018) Deep similarity learning for multimodal medical images (MICCAI)

### Reinforcement learning

Mostly used for *rigid registration*

- Low-resolution transformation model for deformable registration

  Krebs et al. (2017) Robust non-rigid registration through agent-based action learning (MICCAI)

# Where are we in the image registration “recipe”?

**Transformation model:**
(e.g., rigid, affine, deformable)

$T(x', \theta)$

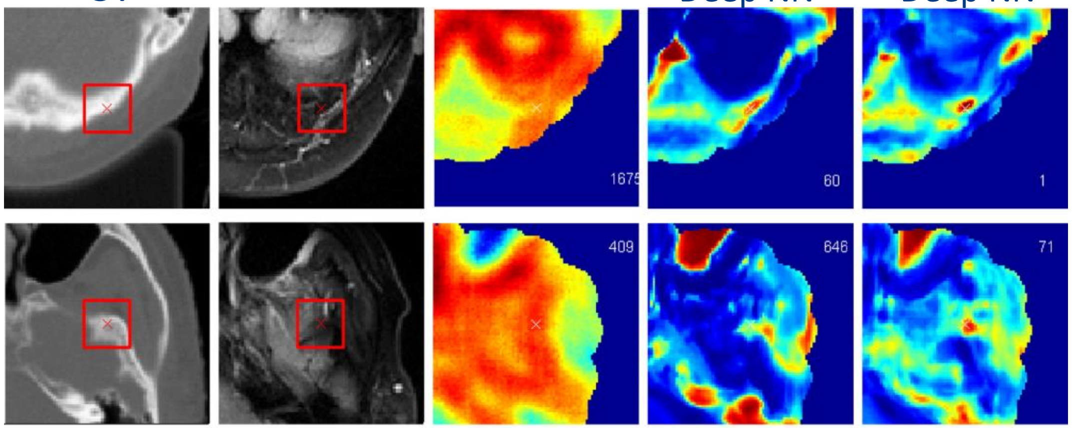

**Similarity measure:**
(e.g. SSD, CC, MI, MSE)

$S(x, x')$

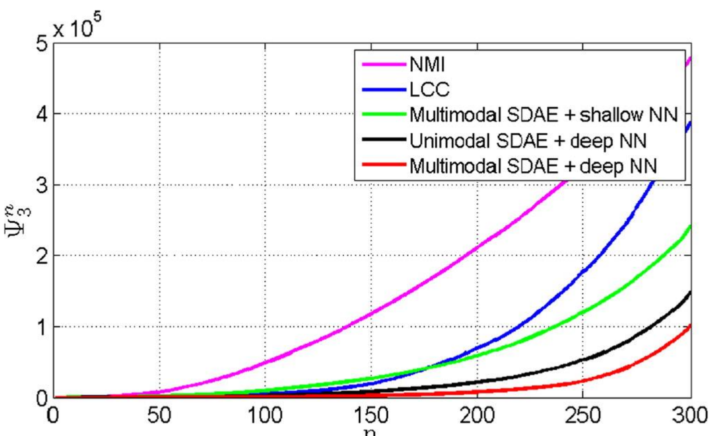
**Optimization:**

$$\hat{\theta} = \max_{\theta} S\big(x, T(x', \theta)\big)$$

# Learned multi-modal feature extraction




Cumulative sum of prediction errors




No. of patches



Cheng, Xi, Li Zhang, and Yefeng Zheng. "Deep similarity learning for multimodal medical images." *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 6.3 (2018): 248-252

# Where are we in the image registration “recipe”?

**Transformation model:**
(e.g., rigid, affine, deformable)

$T(x', \theta)$

Image registration

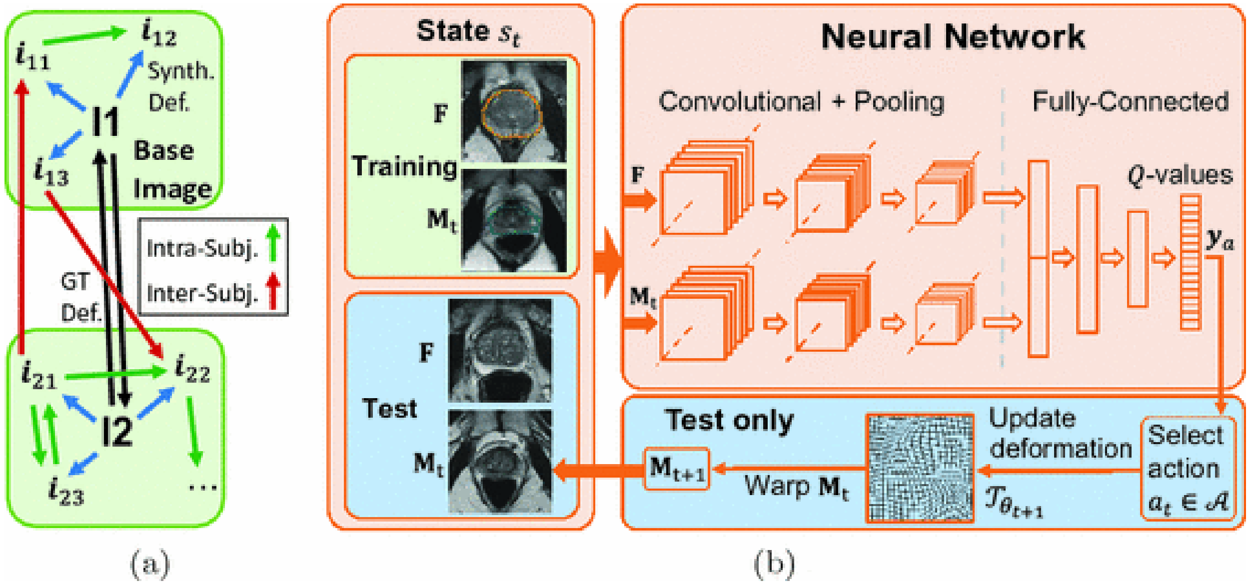
**Similarity measure:**
(e.g. SSD, CC, MI, MSE)

$S(x, x')$

**Optimization:**

$$\hat{\theta} = \max_{\theta} S\big(x, T(x', \theta)\big)$$

# Reinforcement learning for image registration




(a) Training Data Generation: **Synthetic deformations** (blue arrows) and inter-subject GT deformations (black) are used for intra- (green) and inter-subject (red) image pairs for training. (b) Dual-stream network used for Q-value prediction including complete single-stage Markov Decision Process for testing (blue background).



Krebs, Julian, et al. "Robust non-rigid registration through agent-based action learning." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2017.

## B. Supervised methods

Require many known transformations for training, use ground truth labels to calculate the loss

### Fully supervised

2D (FlowNet) vs. 3D approaches (e.g. 3D U-net); mostly patch-based (e.g. $128^3$)

- *Generation of ground truth transformations* → many, e.g. Eppenhof et al. (2018)
  Pulmonary ct registration through supervised learning with convolutional neural networks (IEEE trans. on med. imaging)
- *Large Deformation Diffeomorphic Metric Mapping (LDDMM)* → Yang et al. (2017)
  Quicksilver: Fast Predictive Image Registration - a Deep Learning Approach (NeuroImage)

### Weakly supervised

Use overlap between segmentations or a similarity metric between *M* and *F* combined with ground truth

- MR-US registration
  - *CNN* → Hu et al. (2018) Weakly-supervised convolutional neural networks for multimodal image registration (Medical Image Analysis)
  - *Generative Adversarial Networks (GANs)* → Yan et al. (2018) Adversarial image registration with application for mr and trus image fusion (arXiv preprint: 1804.11024)

Supervised learning of the transformation requires a ground truth!

Questions:

- What is the ground truth of an image registration task?
  - Parametric?
  - Non-parametric?

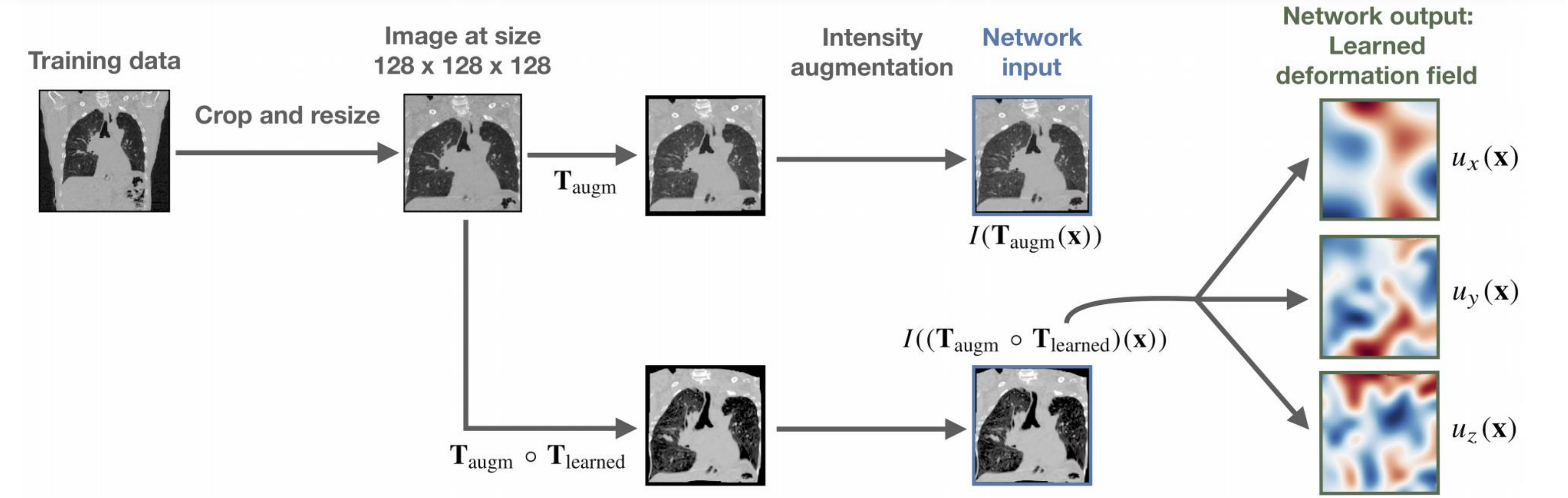
- How can we acquire this?

# Simulation of displacement vector fields

TABLE I

PARAMETERS FOR THE B-SPLINE TRANSFORMATIONS IN THE TRAINING SET. THE GRID SIZE INDICATES THE NUMBER OF GRID POINTS IN EACH DIMENSION. THE DISPLACEMENTS ARE SAMPLED FROM UNIFORM DISTRIBUTIONS IN THE GIVEN RANGES.

| | | **Grid point displacement ranges** (voxels) | | |
|---|---|---|---|---|
| **T** | **Grid size** | $x$ | $y$ | $z$ |
| $\mathbf{T}_{\text{augm}}$ | $2 \times 2 \times 2$ | $[-3.2, 3.2]$ | $[-6.4, 6.4]$ | $[-12.8, 12.8]$ |
| $\mathbf{T}^{\text{coarse}}_{\text{learned}}$ | $4 \times 4 \times 4$ | $[-3.2, 3.2]$ | $[-6.4, 6.4]$ | $[-12.8, 12.8]$ |
| $\mathbf{T}^{\text{fine}}_{\text{learned}}$ | $8 \times 8 \times 8$ | $[-3.2, 3.2]$ | $[-3.2, 3.2]$ | $[-3.2, 3.2]$ |






Eppenhof, Koen AJ, and Josien PW Pluim. "Pulmonary CT registration through supervised learning with convolutional neural networks." *IEEE transactions on medical imaging* 38.5 (2018): 1097-1105.

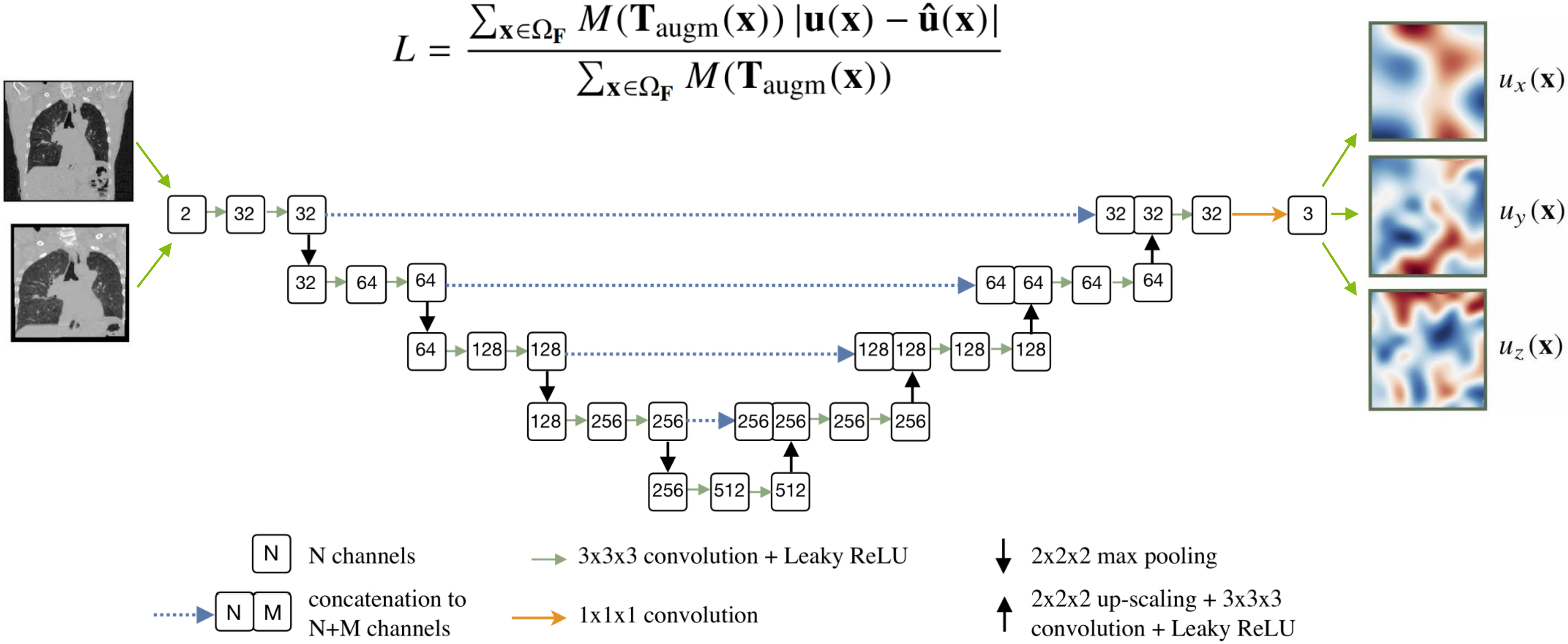
# **Learning the displacement vector field** (Eppenhof & Pluim, 2018)

Binary mask of the lungs

True vector field

Estimated vector field

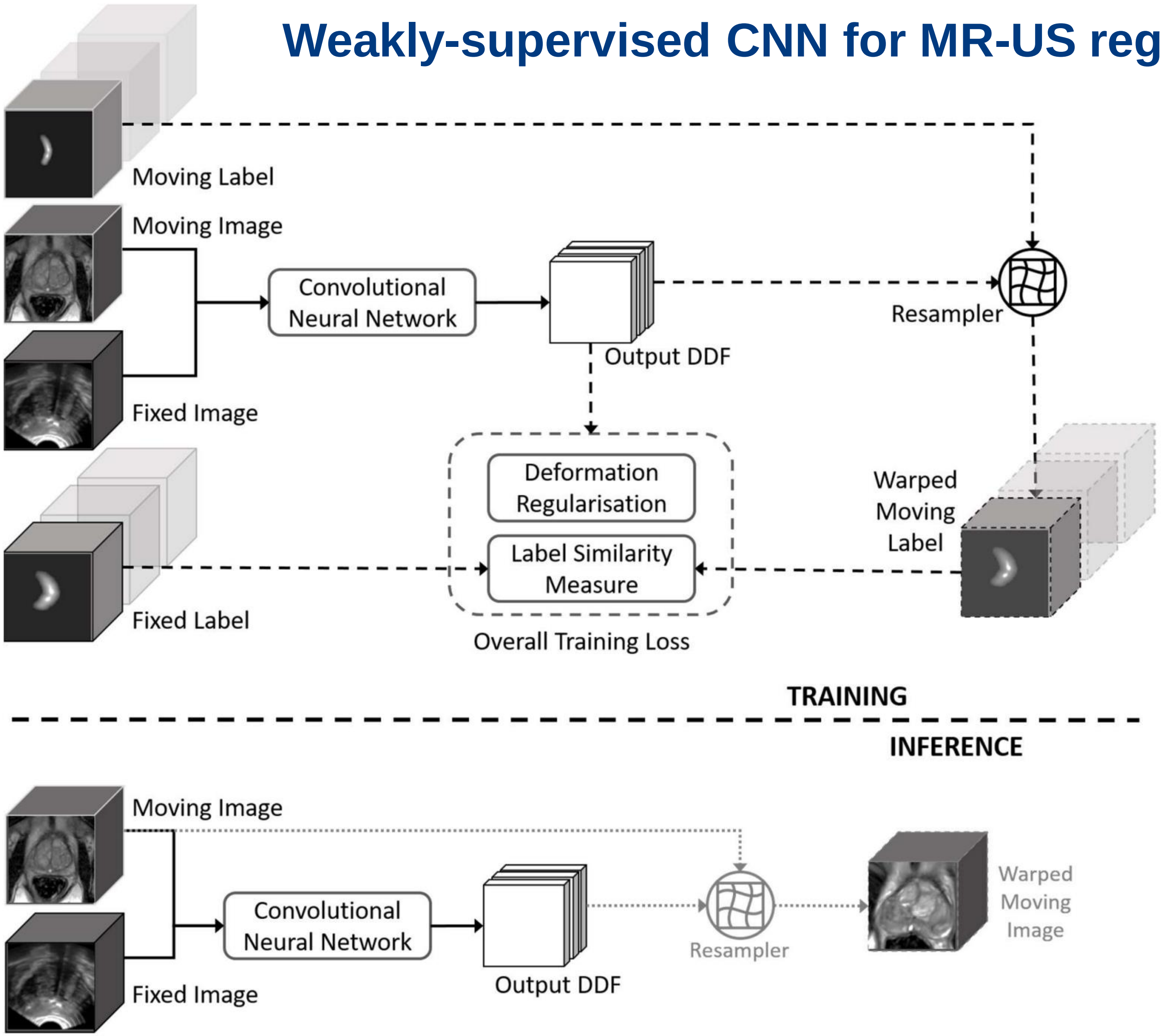
$$L = \frac{\sum_{\mathbf{x}\in\Omega_\mathbf{F}} M(\mathbf{T}_{\text{augm}}(\mathbf{x}))\,|\mathbf{u}(\mathbf{x}) - \hat{\mathbf{u}}(\mathbf{x})|}{\sum_{\mathbf{x}\in\Omega_\mathbf{F}} M(\mathbf{T}_{\text{augm}}(\mathbf{x}))}$$






Eppenhof, Koen AJ, and Josien PW Pluim. "Pulmonary CT registration through supervised learning with convolutional neural networks." *IEEE transactions on medical imaging* 38.5 (2018): 1097-1105.

# Weakly-supervised CNN for MR-US registration




Hu, Yipeng, et al. "Weakly-supervised convolutional neural networks for multimodal image registration." *Medical image analysis* 49 (2018): 1-13.

## C. Unsupervised methods

No ground truth needed. Often use a spatial transformer layer.

**Variational autoencoders** → Krebs et al. (2018) Unsupervised Probabilistic Deformation Modeling for Robust Diffeomorphic Registration. (DLMIA)

**Generative Adversarial Networks (GANs)** → Tanner et al. (2018) Generative Adversarial Networks for MR-CT Deformable Image Registration (CVPR) & Hu et al. (2018) Adversarial Deformation Regularization for Training Image Registration Neural Networks (MICCAI)

**Multi-scale methods:**

- **RegNet** → Sokooti et al. (2017) Nonrigid Image Registation Using Multi-scale 3D Convolutional Neural Networks (MICCAI). NB: dual path
- **ConvNet** → De Vos et al. (2018) A deep learning framework for unsupervised affine and deformable image registration (Medical Image Analysis). NB: chain
- **pgCNN** → Eppenhof et al. (2019) Progressive Growing Convolutional Networks for End-to-End Deformable Image Registration (SPIE medical imaging)

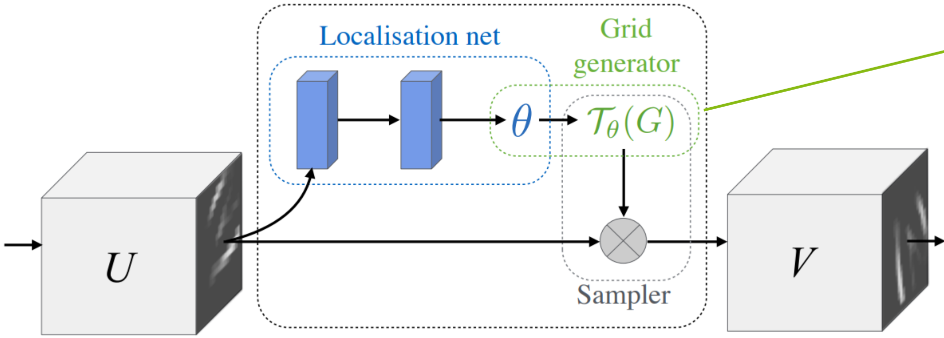
**VoxelMorph (U-net)** → Balakrishnan et al. (2019) VoxelMorph: A Learning Framework for Deformable Medical Image Registration (IEEE tran.med.imaging)

- **Cycle-consistent VoxelMorph** → Kim et al. (2019) Unsupervised Deformable Image Registration Using Cycle-Consistent CNN (MICCAI)

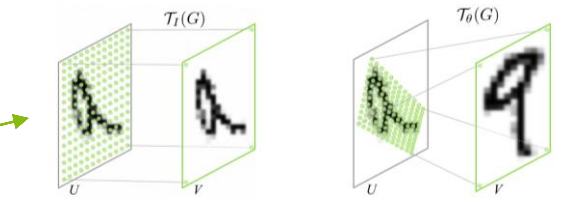
# **Spatial transformer networks** (Jaderberg et al., NIPS 2015)

**Spatial transformer** = a learnable module that explicitly allows the spatial manipulation of data within the network

- *Differentiable*
- Can be inserted into existing convolutional neural networks
- Actively transforms feature maps (conditional on the feature map itself)




Spatial Transformer



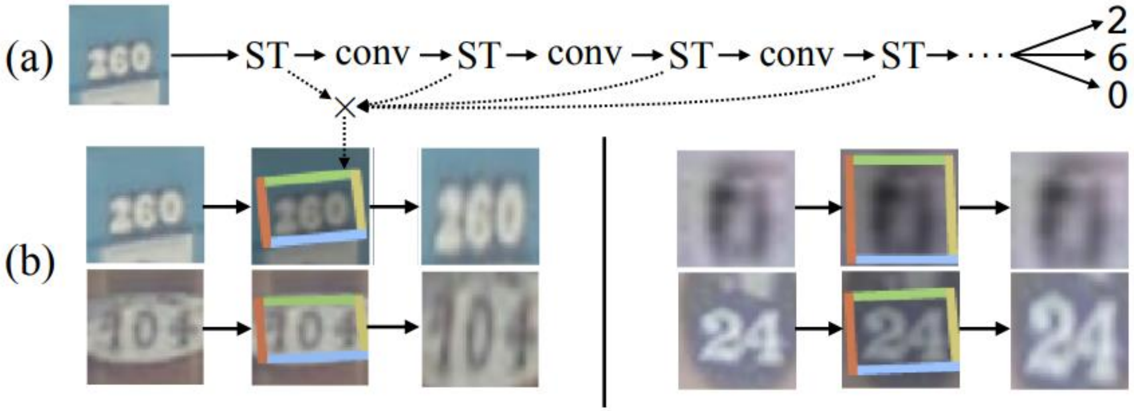

Example $T_\theta(G)$ for an affine transformation:

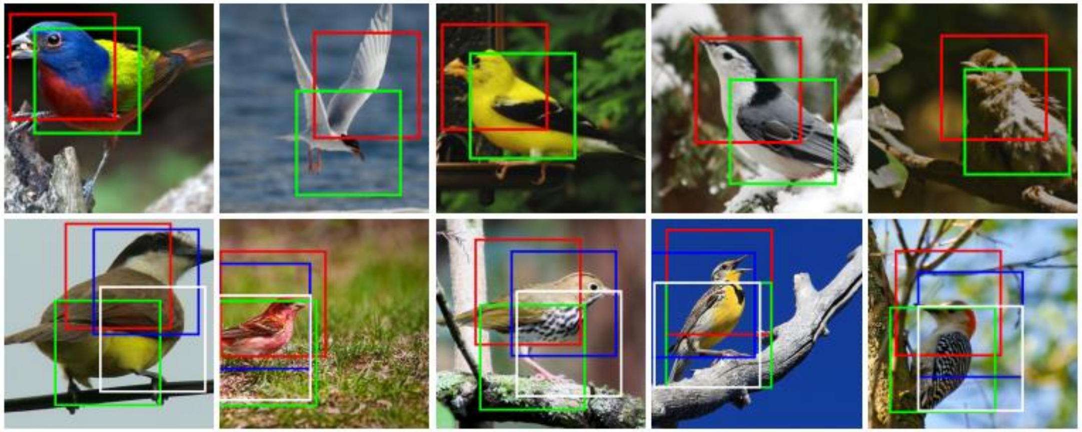
$$\begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = \mathcal{T}_\theta(G_i) = \mathtt{A}_\theta \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{bmatrix} \begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix}$$



Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. "Spatial transformer networks." *Advances in neural information processing systems*. 2015.

## CNNs are in essence not invariant to translation, scale, rotation and more generic warping of the input data



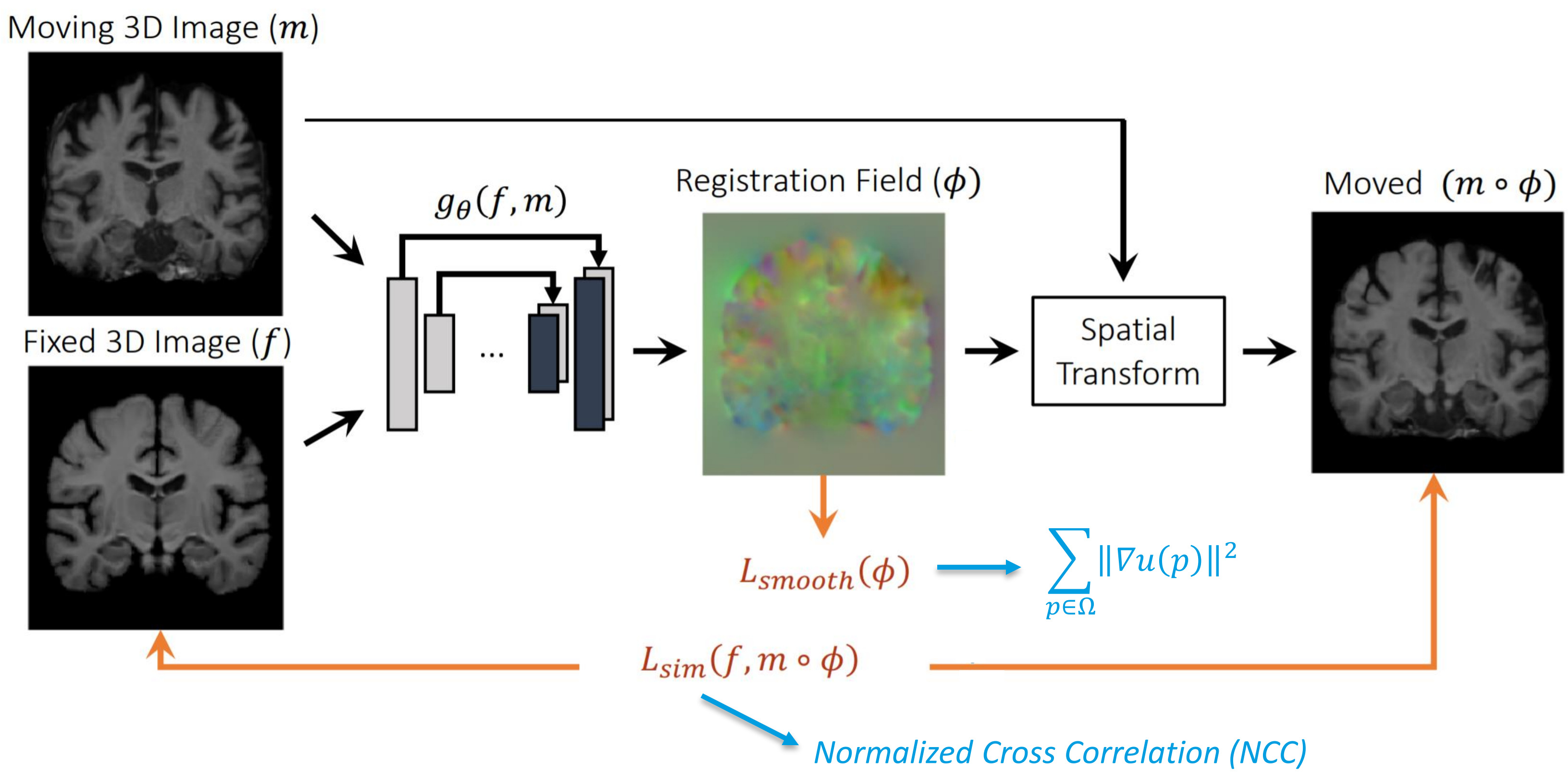





Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. "Spatial transformer networks." *Advances in neural information processing systems*. 2015.

# VoxelMorph

## Example of a popular unsupervised framework for deformable image registration



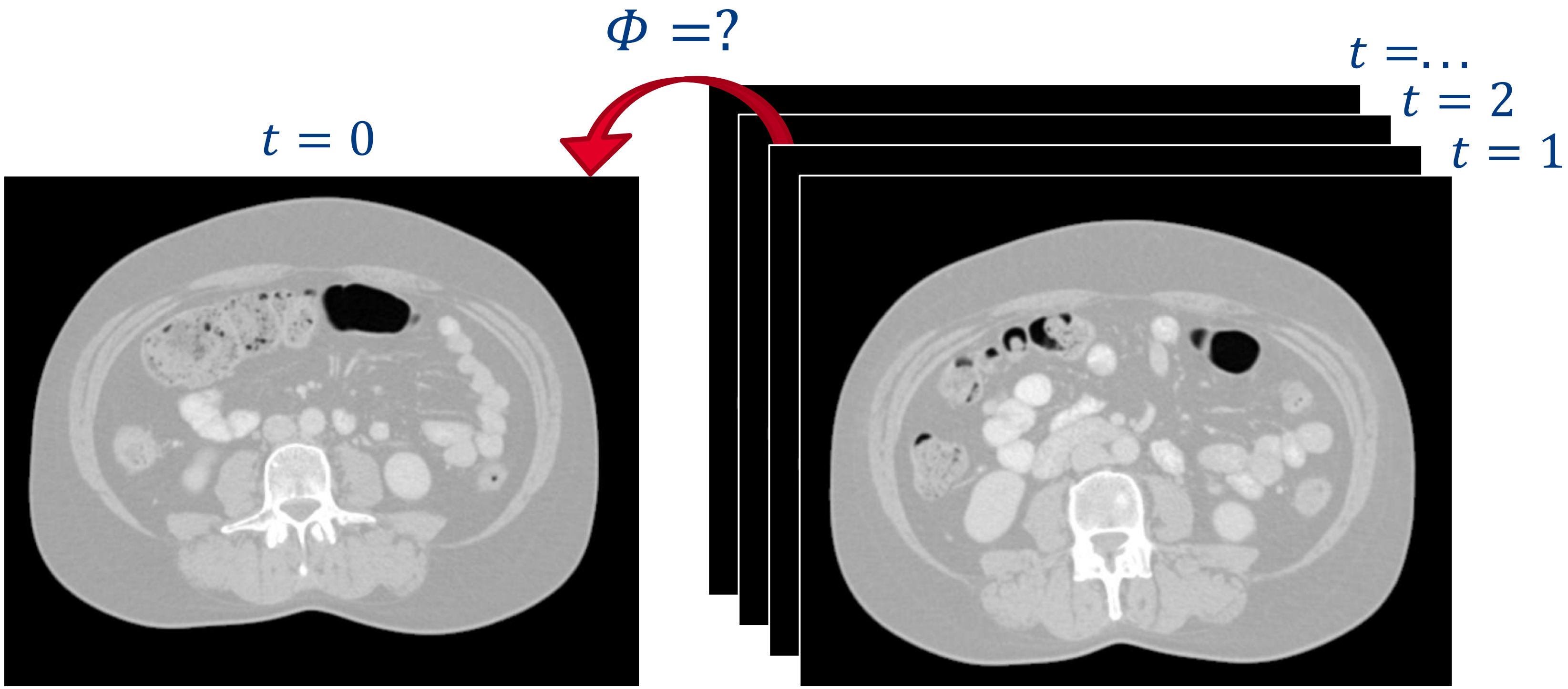

$$\hat{\Phi} = \underset{\Phi}{\operatorname{argmin}}\, L_{sim}(f, m \circ \Phi) + \lambda L_{smooth}(\Phi)$$



Balakrishnan et al. (2018): VoxelMorph: A Learning Framework for Deformable Medical Image Registration. arXiv:1909.05231v3

## Example from my own research:

Deformable registration of abdominal CT scans – a longitudinal dataset

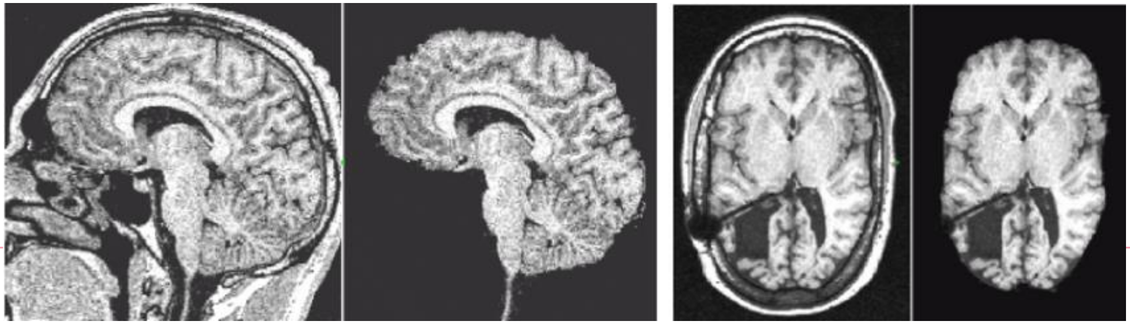



**Resolution**: 512 x 512 x 296 ; **Voxel dimensions**: 0.7 mm x 0.7 mm x 2 mm

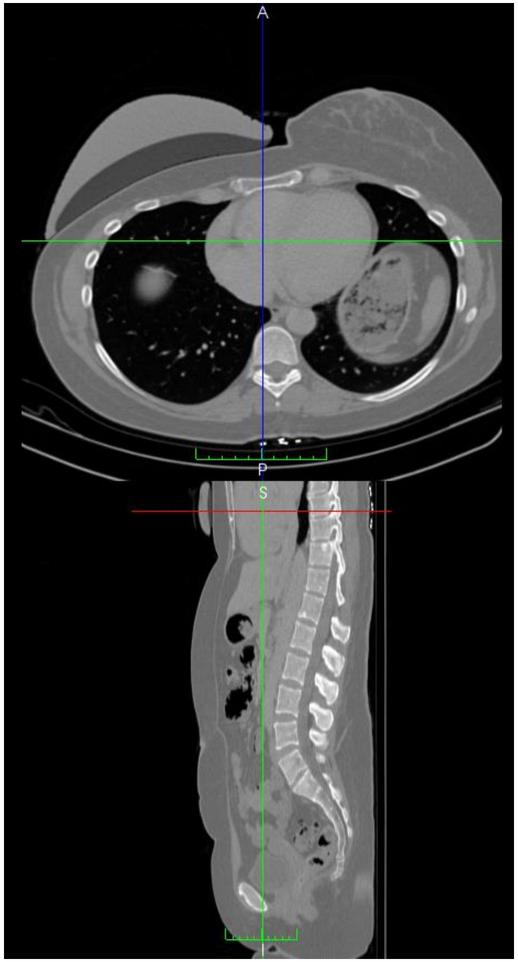
## Dataset description

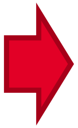
- 140 abdominal and thoracic CT scans of 12 patients with metastasis in the spine
- Gold standard segmentations of vertebrae
- Image registration necessary for reliable response assessment over **several time points**

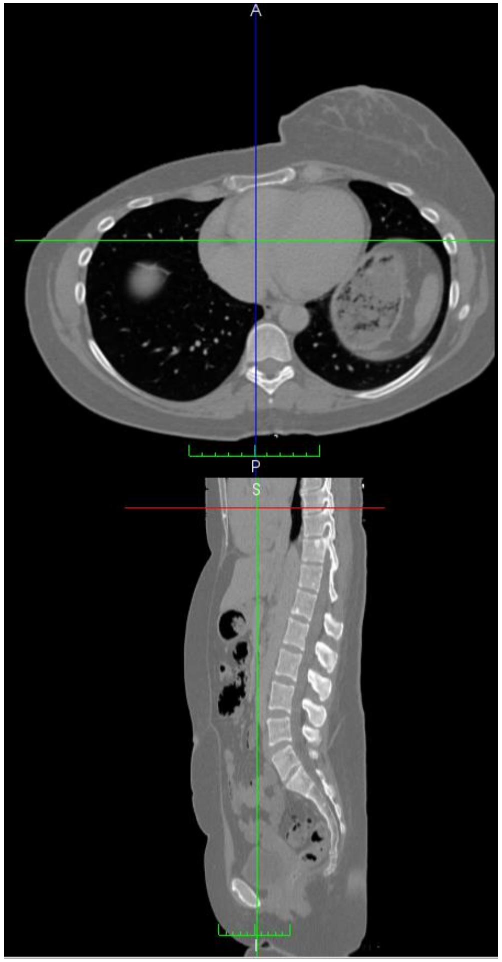
- Brain MR image registration relies on *skull-stripping techniques*
- We developed an automatic region growing algorithm to:
  - Remove patient table
  - Remove non-biological parts (e.g., breast prostheses or clothes) and metal artefacts



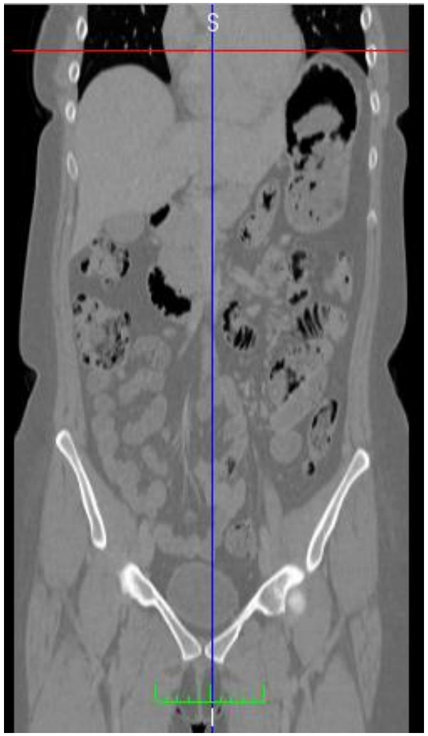
# Region growing – example of CT table and breast prosthesis removal







**DET 183 – Baseline**
(abdominal scan)



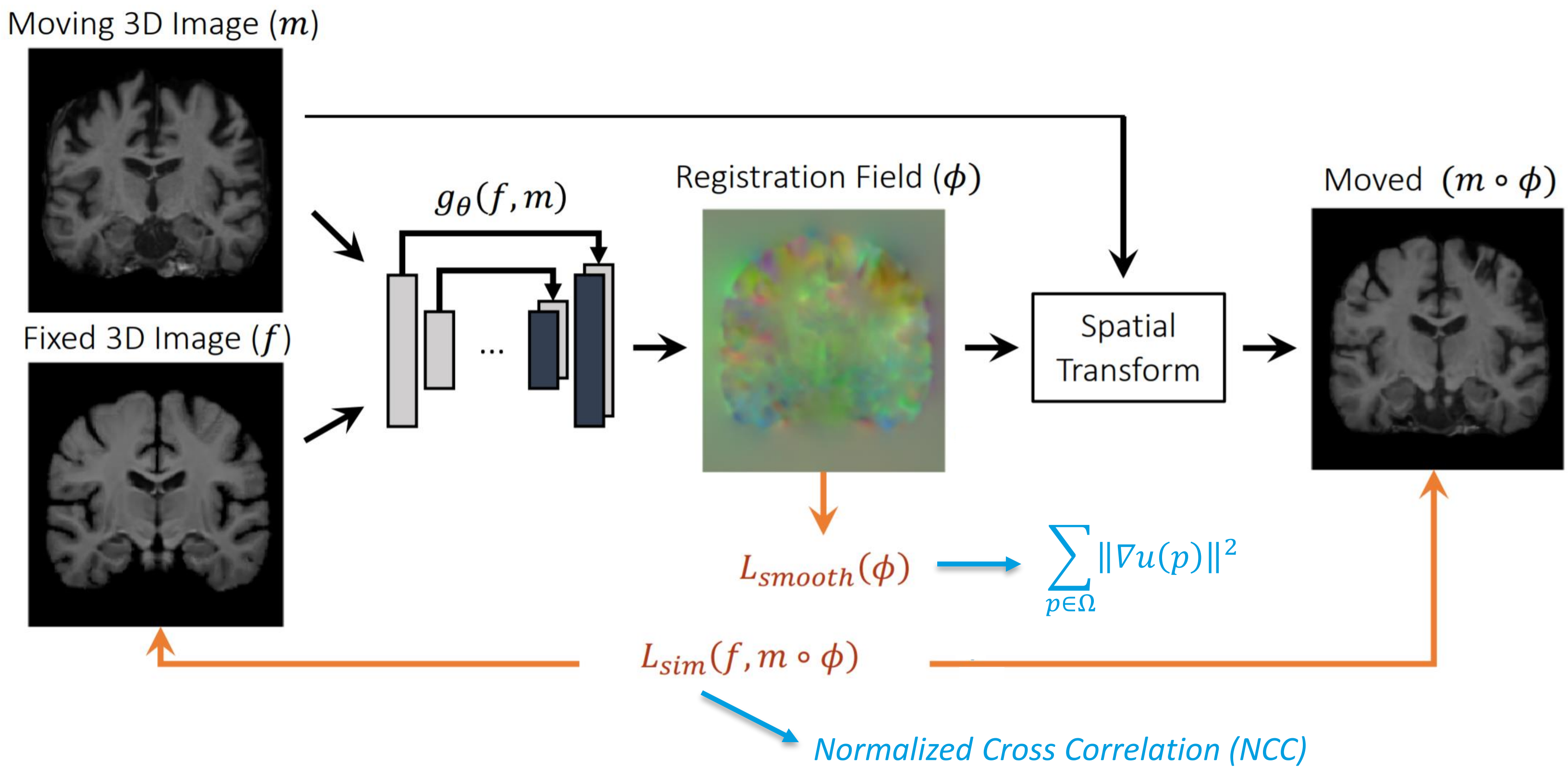
# Pre-processing: resampling and affine pre-alignment

*Fixed* *Moving* *Difference*

**Resolution**: 160 x 160 x 256 ; **Voxel dimensions**: 1 mm x 1 mm x 1 mm

# VoxelMorph

## Example of a popular unsupervised framework for deformable image registration




$$\hat{\Phi} = \underset{\Phi}{\operatorname{argmin}}\, L_{sim}(f, m \circ \Phi) + \lambda L_{smooth}(\Phi)$$



Balakrishnan et al. (2018): VoxelMorph: A Learning Framework for Deformable Medical Image Registration. arXiv:1909.05231v3

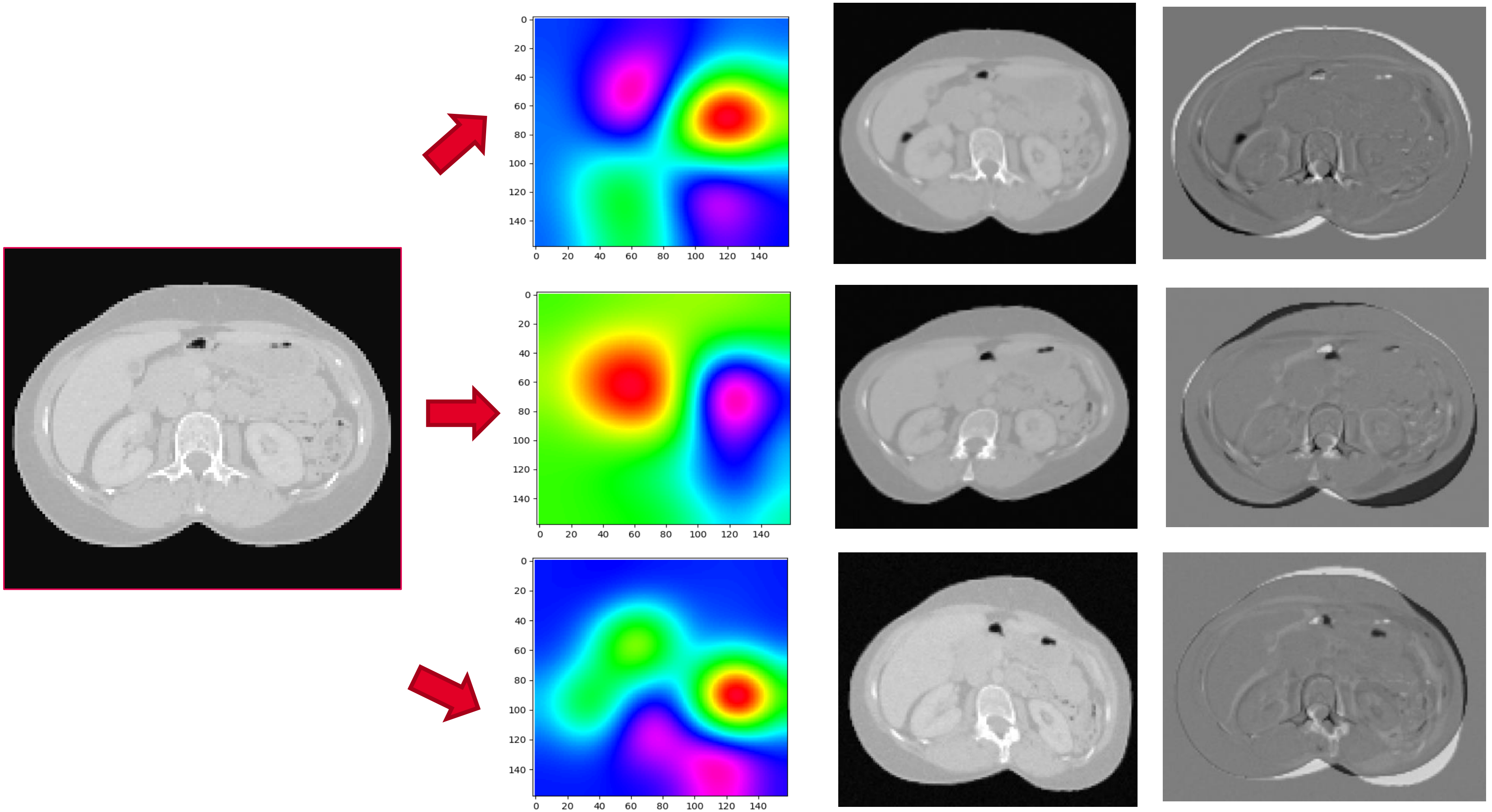
**Research setting** (original VoxelMorph paper):
~6000 brain MRI scans

**Real, clinical setting** (Addenbrooke's hospital in Cambridge, UK):
~140 abdominal CT scans

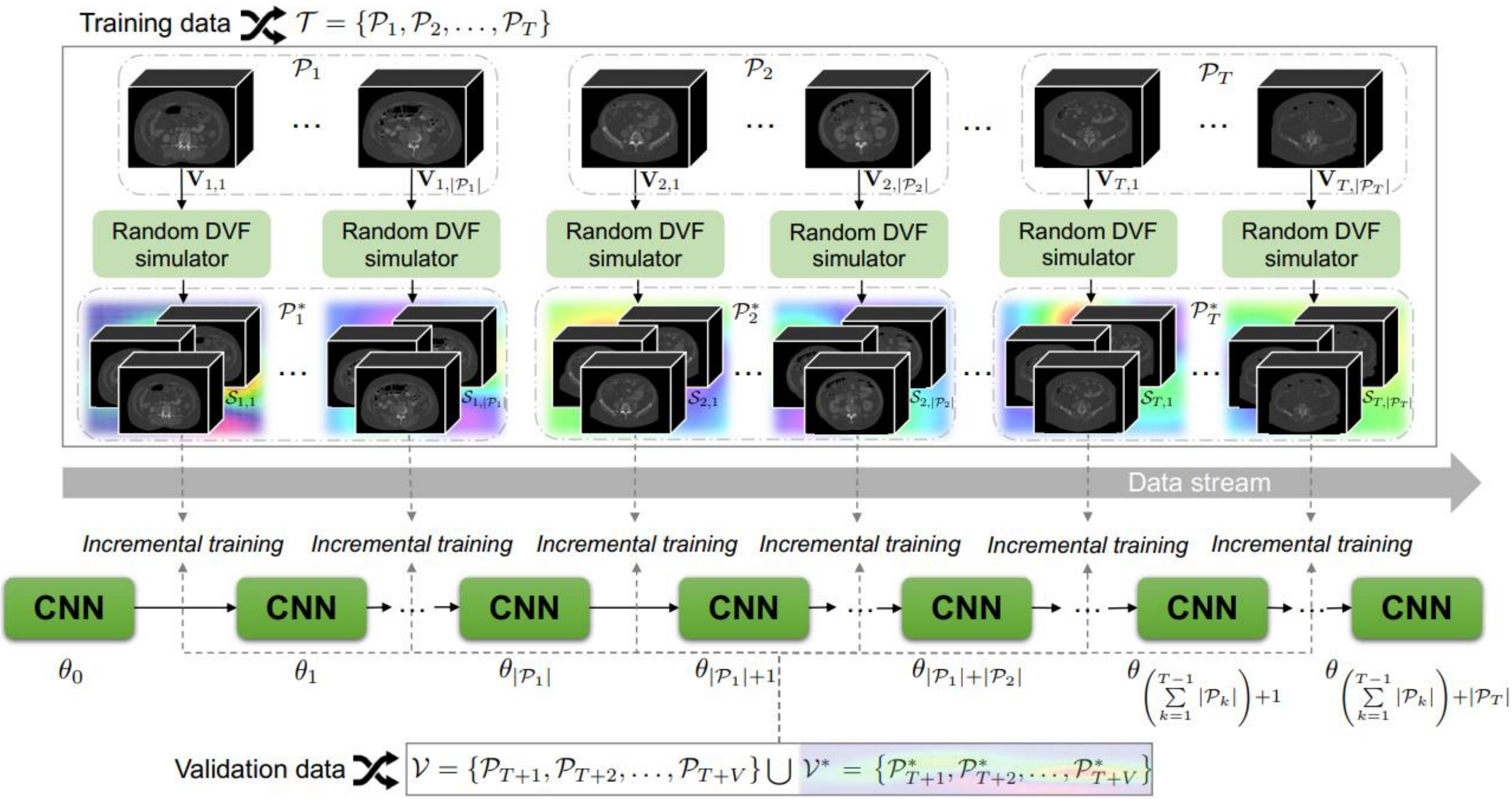
Not enough to train a 3D CNN!

What could we do?

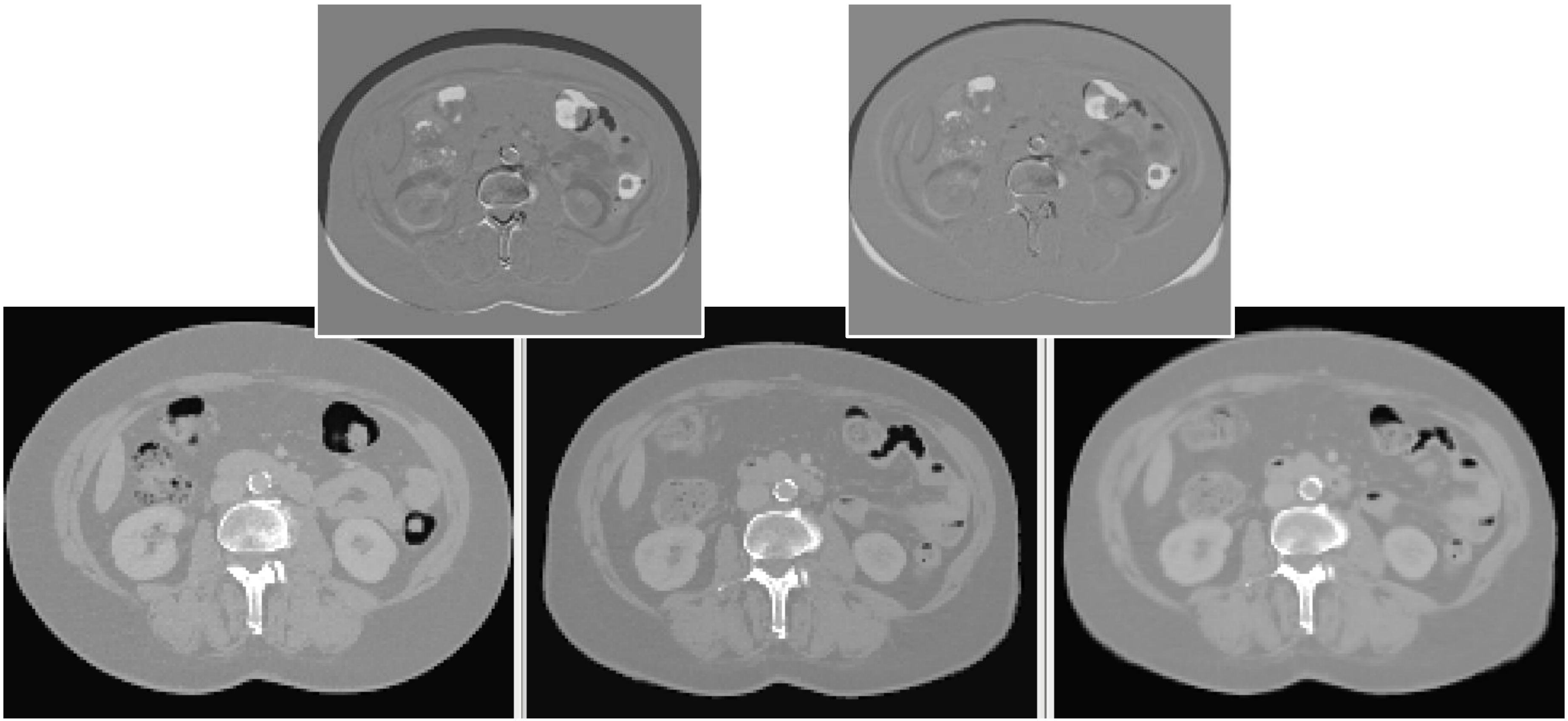
# Displacement vector field simulations



# A novel, incremental training strategy (no exam material)

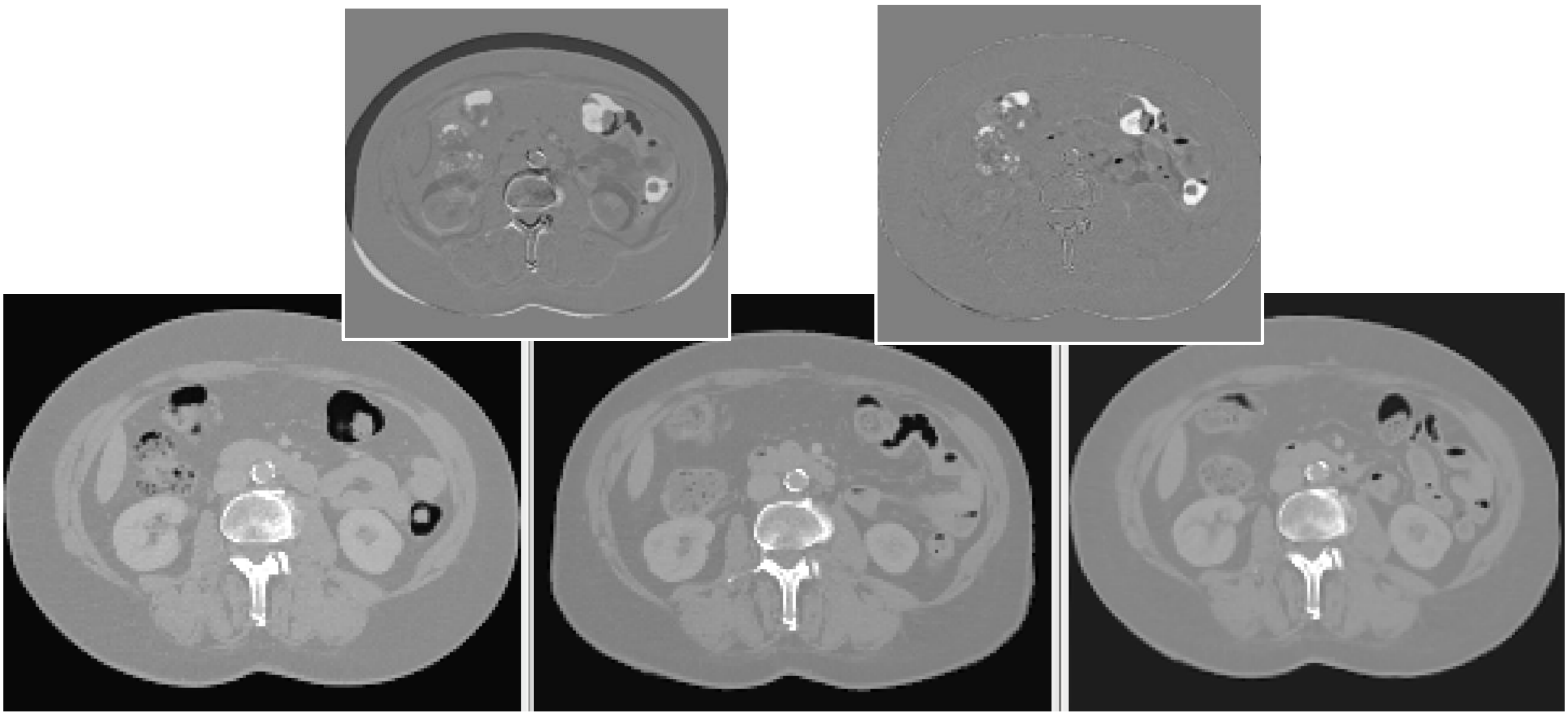


# Results: **VoxelMorph**



Fixed Moving Voxelmorph registration

# Results: the **NiftyReg** toolbox (non-deep learning-based, slow but accurate)



Fixed Moving NiftyReg registration

NiftyReg: an open-source software for efficient medical image registration
http://cmictig.cs.ucl.ac.uk/wiki/index.php/NiftyReg

## Summary

- Parametric vs. non-parametric image registration
- Displacement vector fields
- Formulation of deformable image registration
- Different ways to use deep learning for image registration
  - Deep iterative registration
  - Supervised learning
  - Unsupervised learning
- Disadvantages of deep learning for image registration
  (performance, # of training data, ground truth, …)

# Feedback Project 1

## General feedback Project 1

- **Results**: make sure to organize the results well, logical order, caption for each figure!
- **Discussion**: do the discussed topics directly follow from the results?
- **Use references**! (NB: scientific writing)
- **Reproducibility**: e.g. create a notebook that allows others to easily repeat the experiment or run the code

Grades and group-specific feedback have been communicated with all groups via Canvas, if have questions about this, ask your TA (or me).

# Preparing for the exam

**How can I prepare for the exam?**

- (Re)study all exercises and PCA demo (NB: open questions)
  - Answers on Canvas (soon)
- Questions in the lecture slides
- Discussion points CAD by Mitko on Canvas
- Group exercises lecture 2 by Friso (“Deep learning frameworks & applications”)
- Some example exam questions on next slides

## Question 1

The sequence of transformations that registers a moving to a fixed image is the following: scaling by 0.5 in the x-direction and by 2 in the y-direction, followed by counterclockwise rotation by π/4 radians.

Which of the following sequence of transformation matrices should be applied to the grid G of the output (transformed) image when performing inverse mapping? Motive your answer.

(a)

$$\begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix} \mathbf{G}$$

(b)

$$\begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix} \begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix} \mathbf{G}$$

(c)

$$\begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} \cos(-\frac{\pi}{4}) & -\sin(-\frac{\pi}{4}) \\ \sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{bmatrix} \mathbf{G}$$

(d)

$$\begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \cos(-\frac{\pi}{4}) & -\sin(-\frac{\pi}{4}) \\ \sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{bmatrix} \mathbf{G}$$

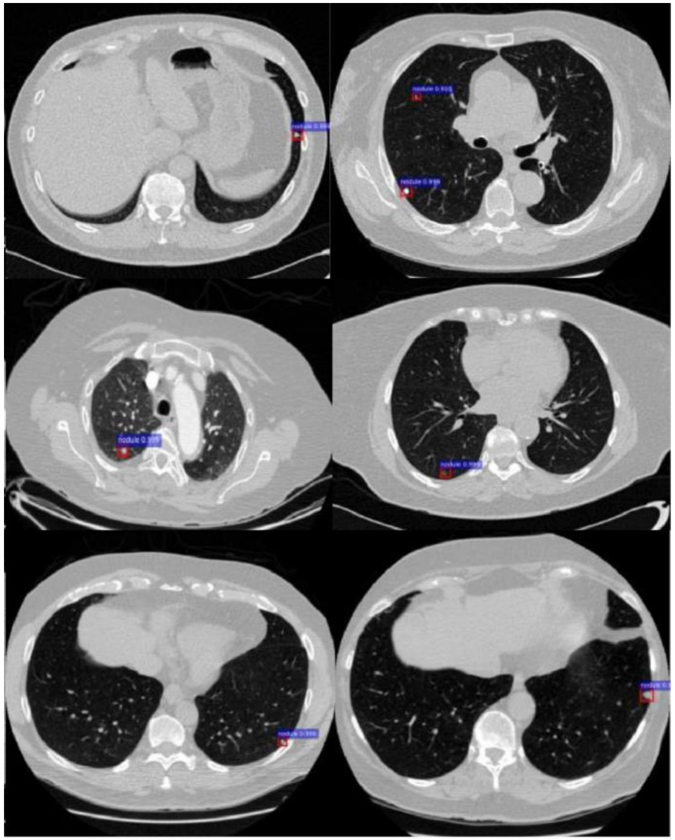
# [explanation]

## Question 2

You are given a dataset of 3D MRI scans of the brain. For each scan, you are also given the manually measured brain surface area in $cm^2$. You want to use this dataset to train a machine learning model that given a 3D MRI scan as an input will output a prediction of the brain surface area. Is this a classification or a regression problem? Motivate your answer.

## Question 3

You are asked to develop a convolutional neural network to automatically detect small lung nodules in CT images (see image on the right). The available dataset consists of medical CT images and ground truth segmentations of the nodules delineated by three experienced radiologists. When you visualize the output of the network, the results look promising.



a) Explain how you would evaluate the performance of your algorithm (use “training set” and “validation set” in your answer).

b) Which quantitative error metric(s) would you choose to report, and why?

c) Can you identify factors that influence the performance of your algorithm? Name three.

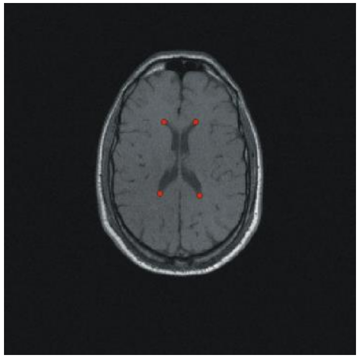


Image source: Xie, Hongtao, et al. "Automated pulmonary nodule detection in CT images using deep convolutional neural networks." *Pattern Recognition* 85 (2019): 109-119.

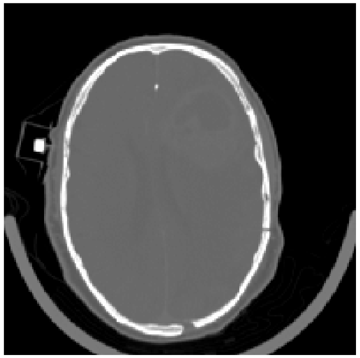
## Some examples of open questions:

- Name two applications of image registration in radiotherapy. Give a short description of every use.
- What type of markers do you see in image (a)? What about image (b)? Give an advantage and disadvantage of using such markers for medical image registration.

(a)



(b)



- What is the difference between linear and logistic regression?
- What is meant with a receptive field?

- Active shape models → your turn! ☺