Imperial College London

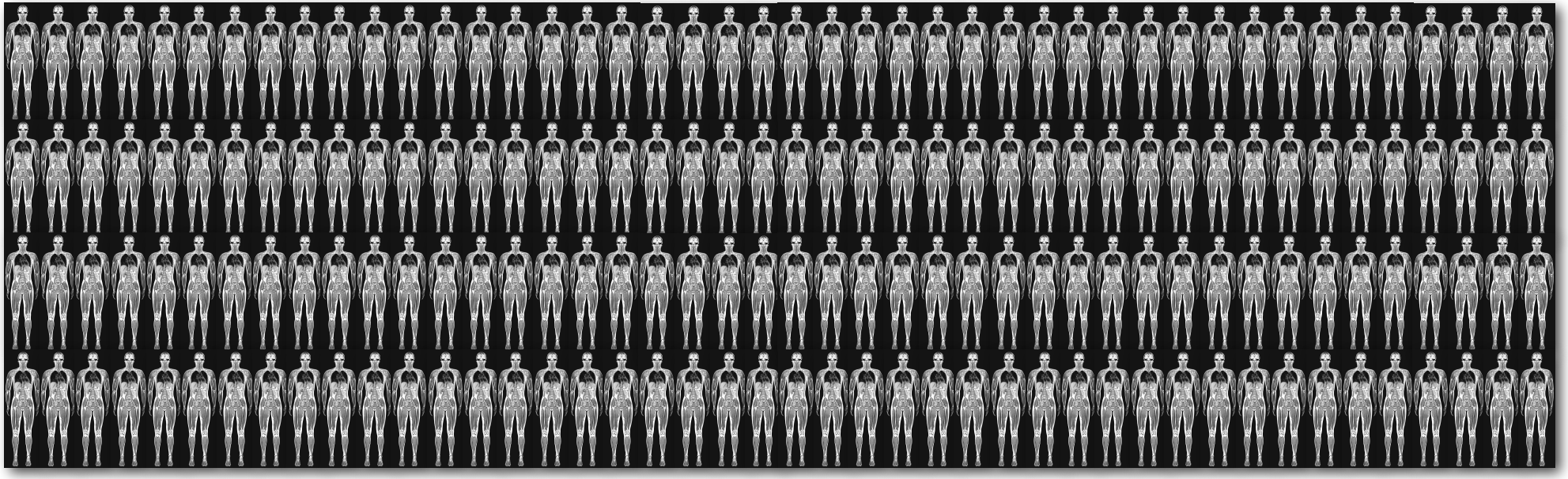
Deep Learning for medical image reconstruction, super-resolution, classification and segmentation

Daniel Rueckert
Biomedical Image Analysis Group
Department of Computing
Imperial College London, UK

Machine learning for medical imaging: Opportunities

- Big data is slowly arriving in medical imaging

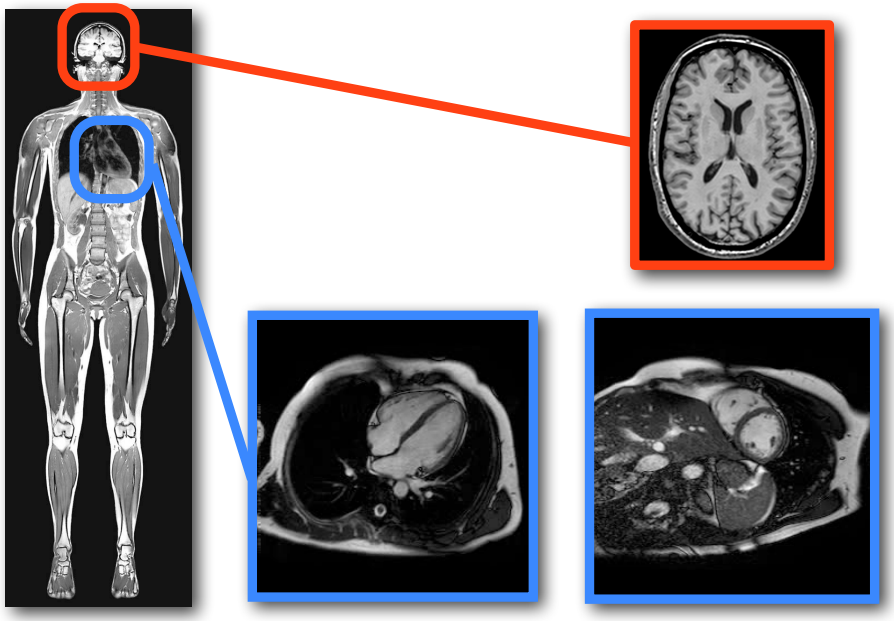
UK Biobank will provide large-scale imaging data from 100,000 subjects



Machine learning for medical imaging: Opportunities

- Big data is slowly arriving in medical imaging

UK Biobank will provide large-scale imaging data from 100,000 subjects



+

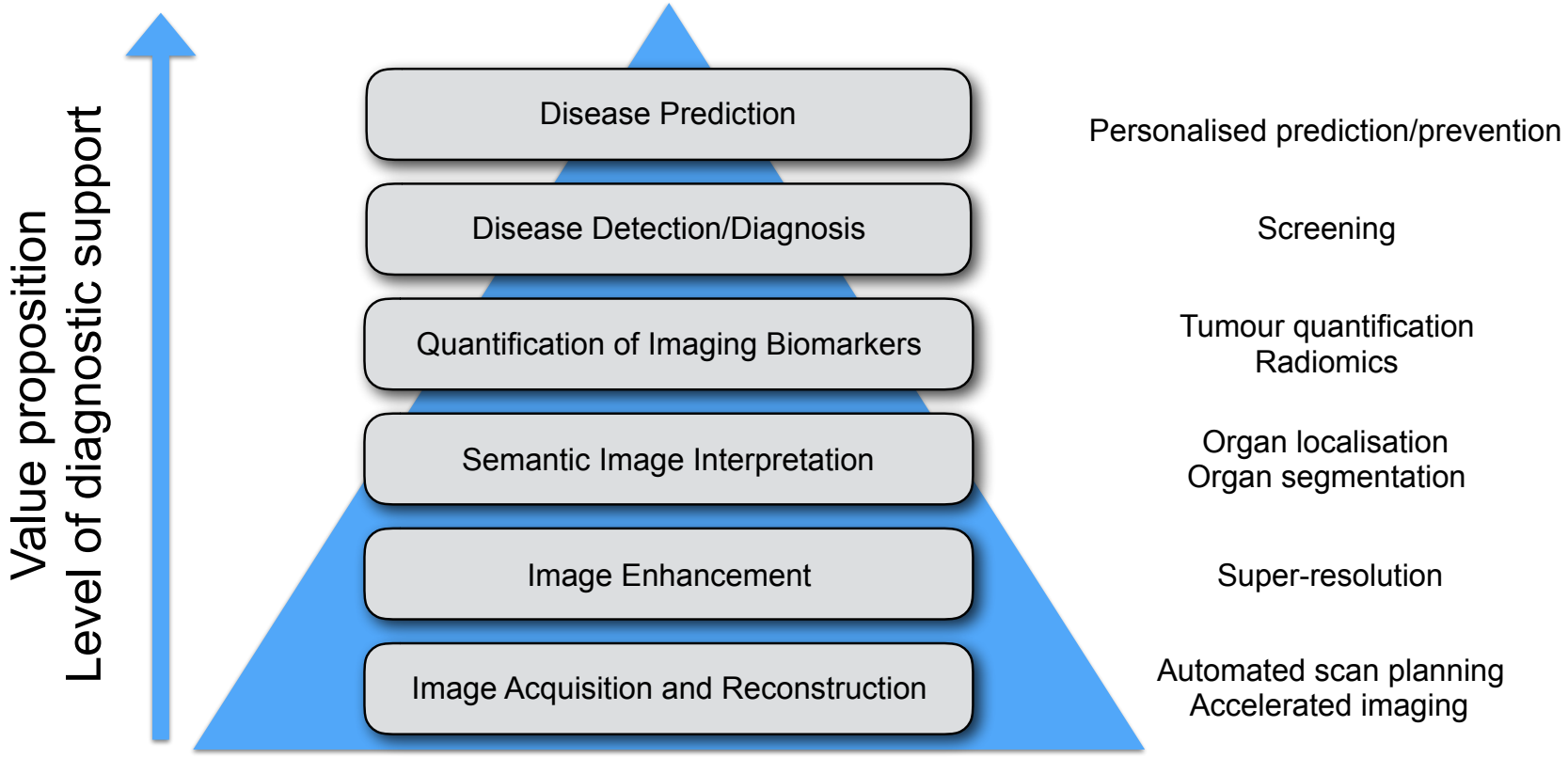
Lifestyle

Genetics

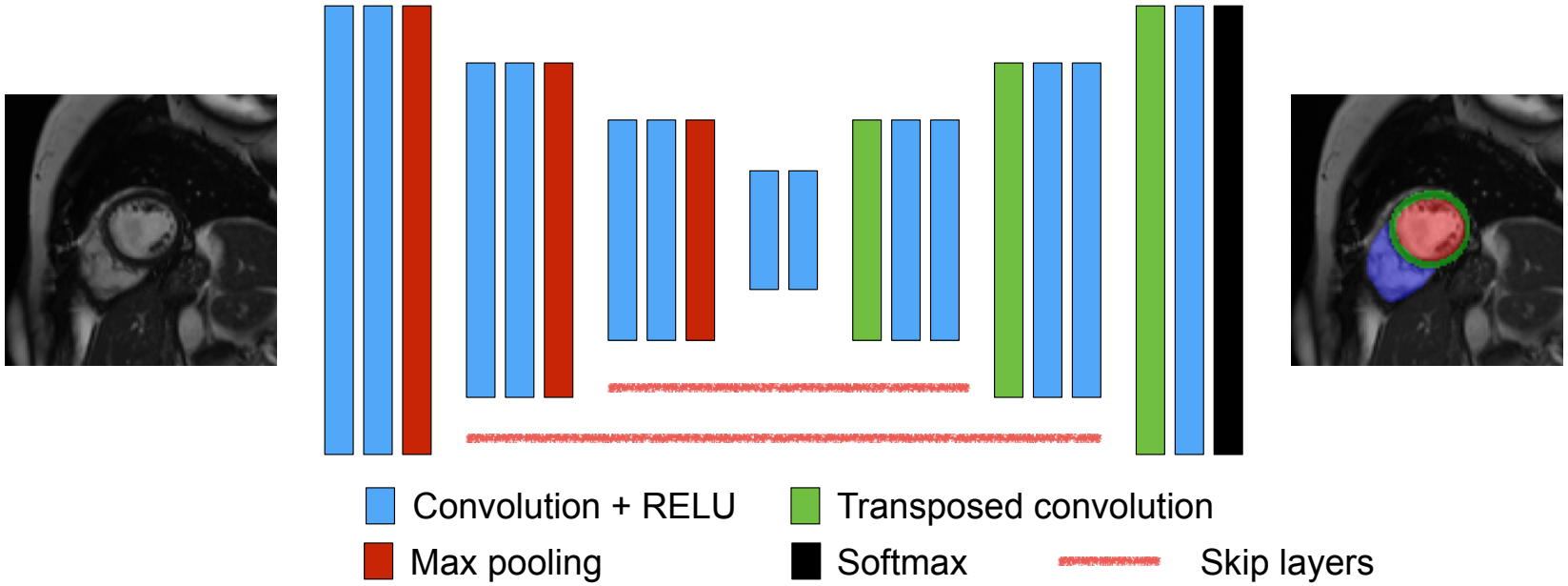
Clinical records

Machine learning for medical imaging: Opportunities

- Machine learning techniques are starting to reach levels of human performance in challenging visual tasks

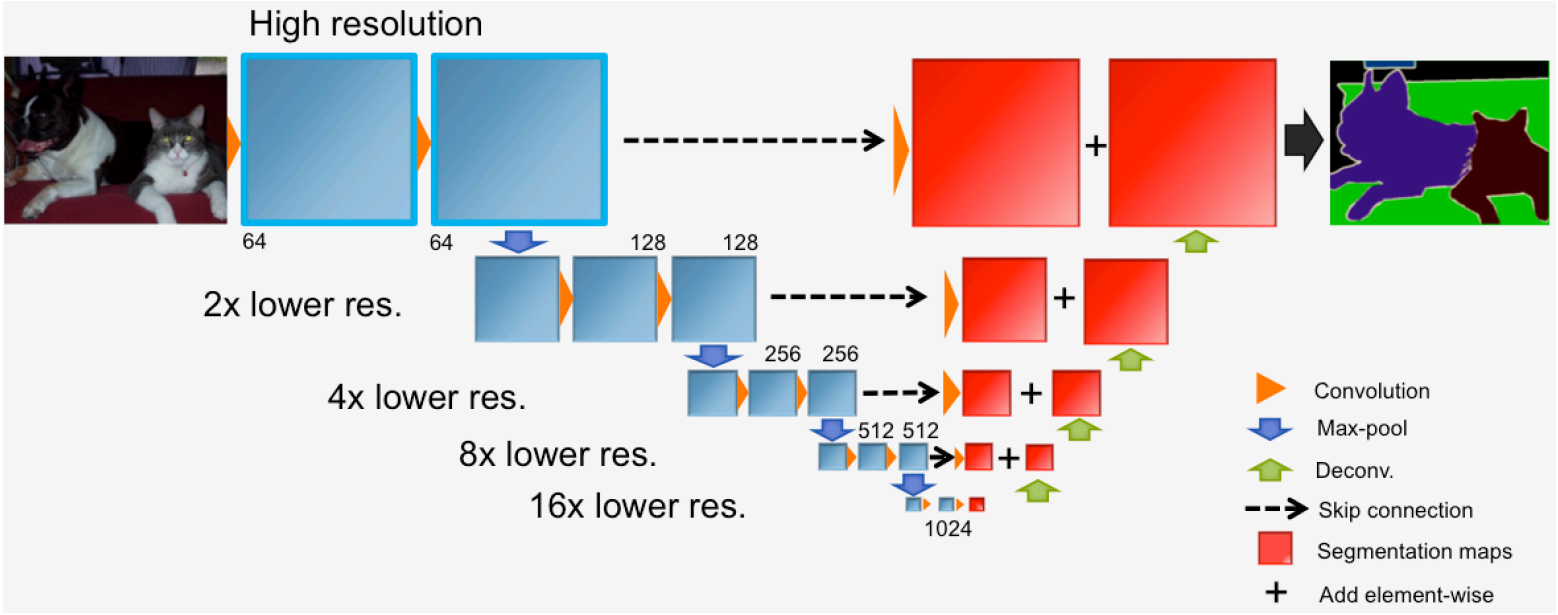


Recall: CNNs: Encoder – Decoder networks



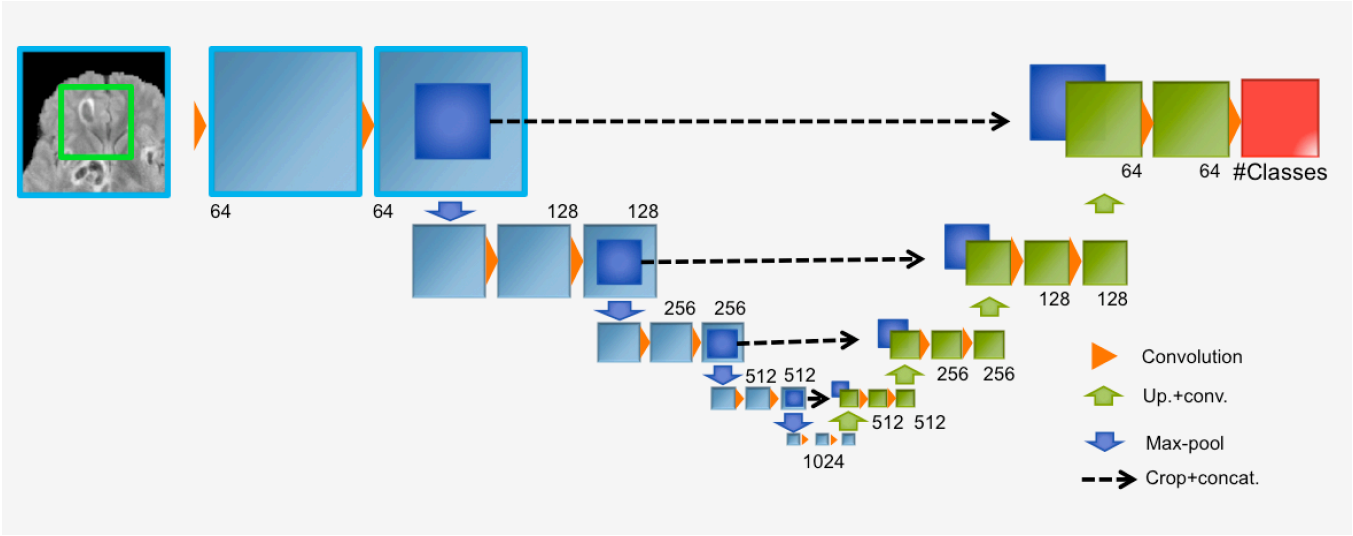
CNNs: Encoder – Decoder networks
Fully convolutional network (FCN)

- Long et al., *Fully convolutional networks for semantic segmentation*, CVPR 2015
  - Deconvolution to upsample segmentation.
  - Skip connections to combine multi-scale info for better accuracy.
  - Popularized fully conv net (previously [LeCun '98, Farabet '13])



CNNs: Encoder – Decoder networks
U-Net

- Ronneberger et al., *U-net: Convolutional networks for biomedical image segmentation*, MICCAI'15.
  - Extended FCN decoder with more filters.
  - Learns to up-sample and combine multi-scale features, not segmentations (unlike FCN).
  - Original version uses only "valid" convolution (no padding), processing only real content.
  - Often works better than FCN, with the trade-off of more expensive decoder.



Overview

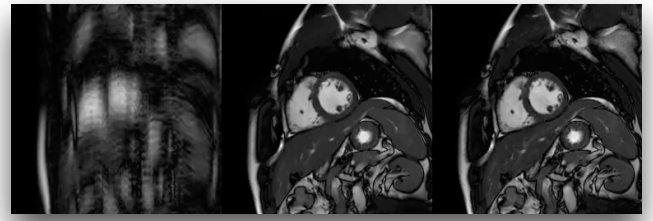


Image reconstruction

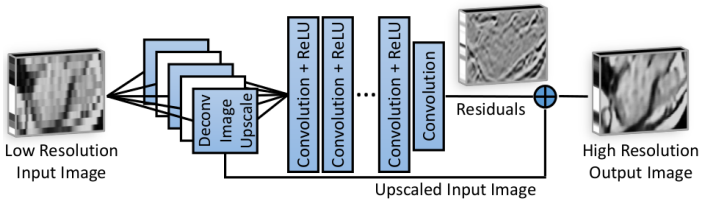


Image super-resolution

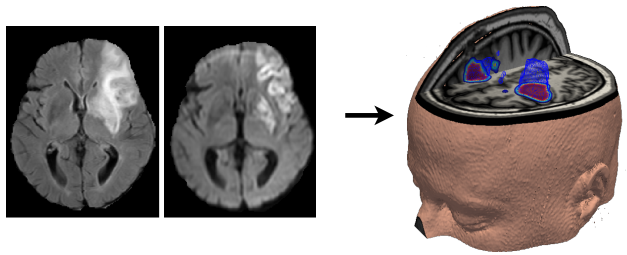


Image segmentation

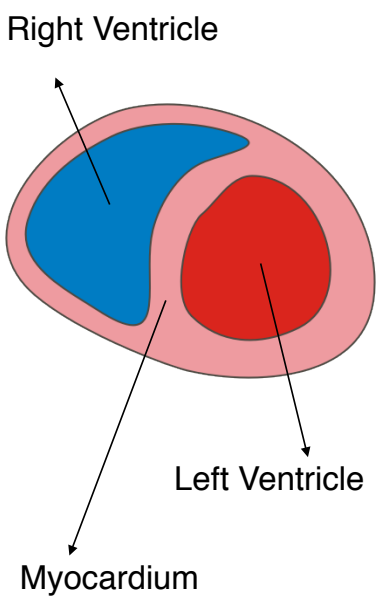
Abdominal View
Confidence: 98%

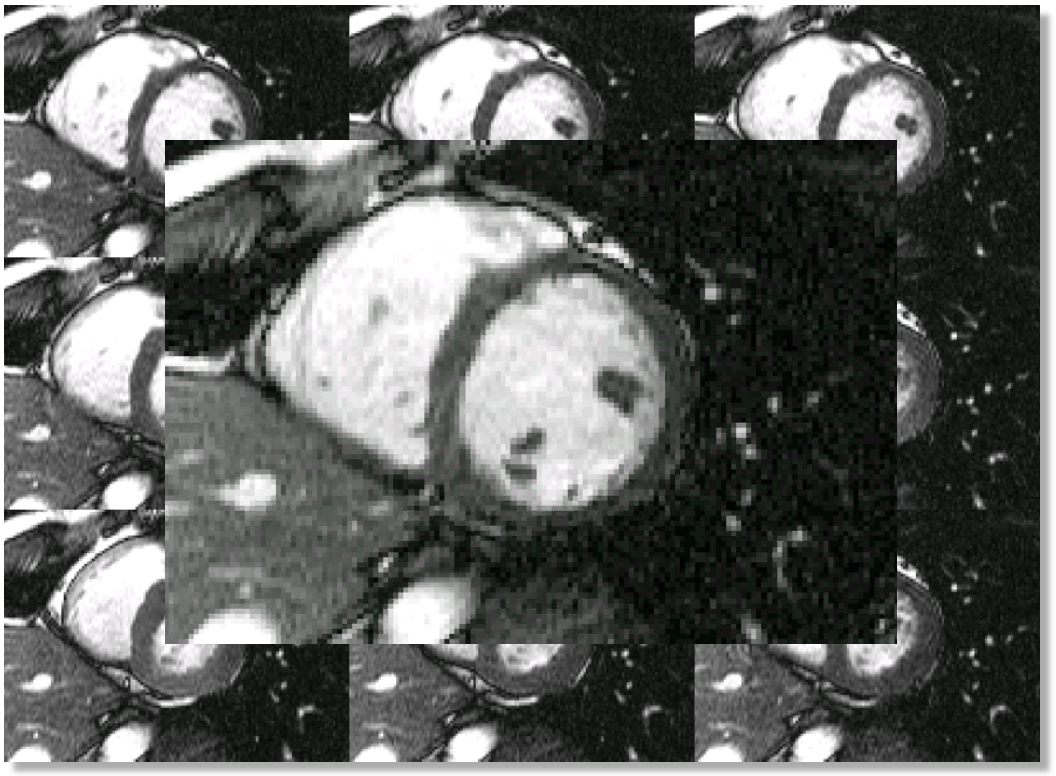
Image classification

MR image acquisition: Challenges

- Magnetic Resonance Imaging (MRI)
  - MRI acquisition is inherently a slow process
  - Slow acquisition is
    - ok for static objects (e.g. brain, bones, etc)
    - problematic for moving objects (e.g. heart, liver, fetus)
  - Options for MRI acquisition:
    - real-time MRI: fast, but 2D and relatively poor image quality
    - gated MRI: fine for period motion, e.g. respiration or cardiac motion but requires gating (ECG or navigators) leading to long acquisition times (30-90 min).

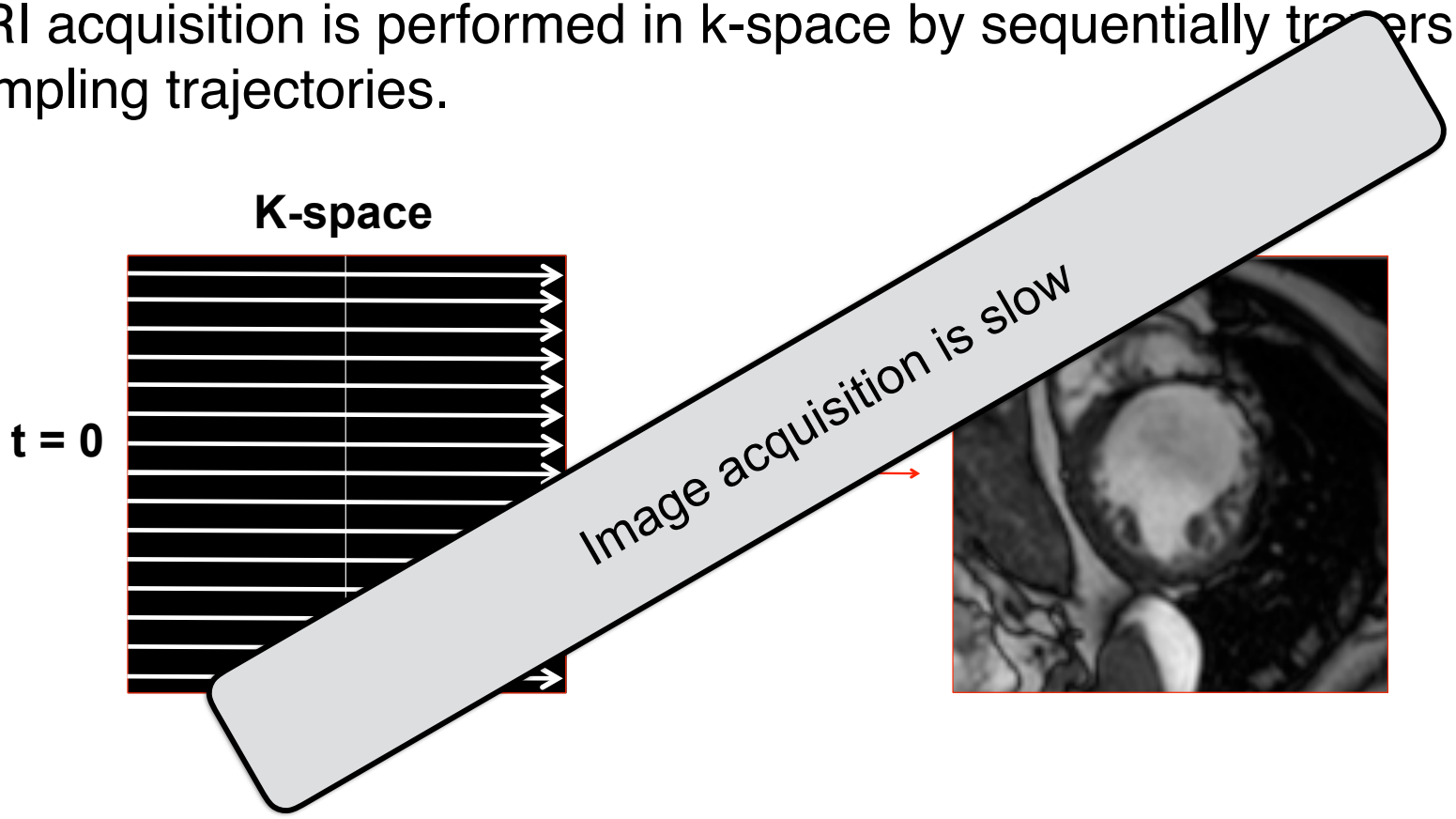
Example: Cardiac imaging





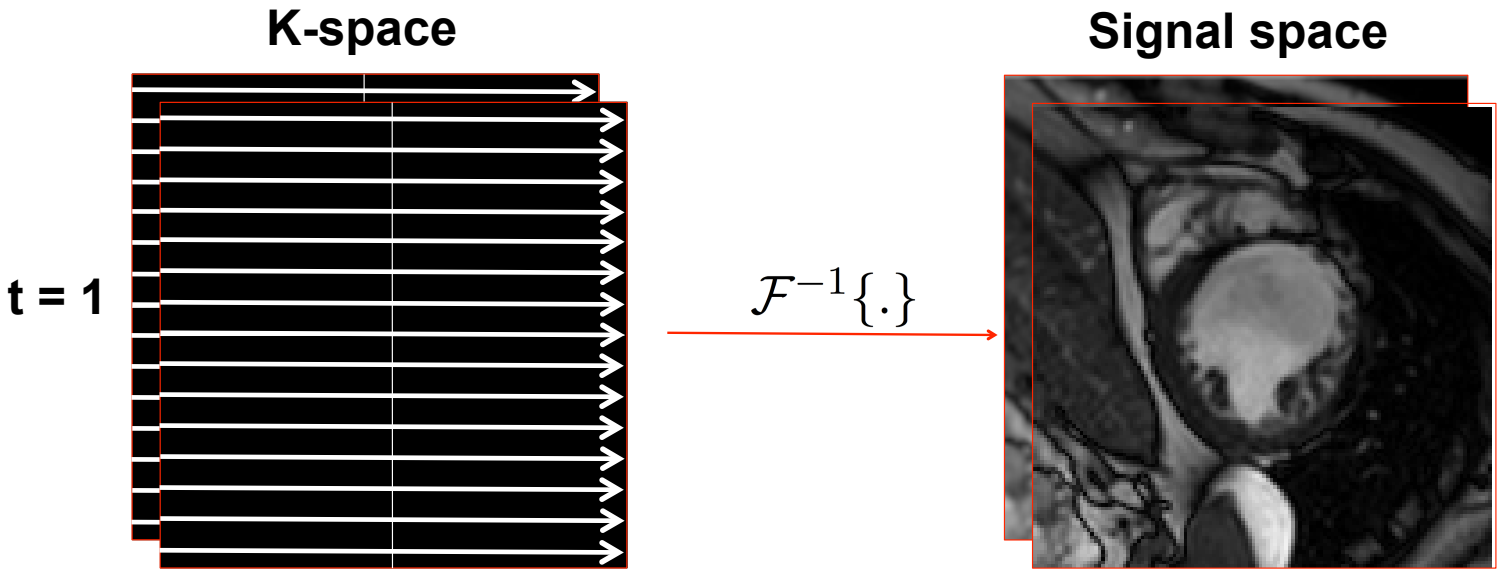
Cardiac MRI: Full acquisition is slow

- MRI acquisition is performed in k-space by sequentially traversing sampling trajectories.



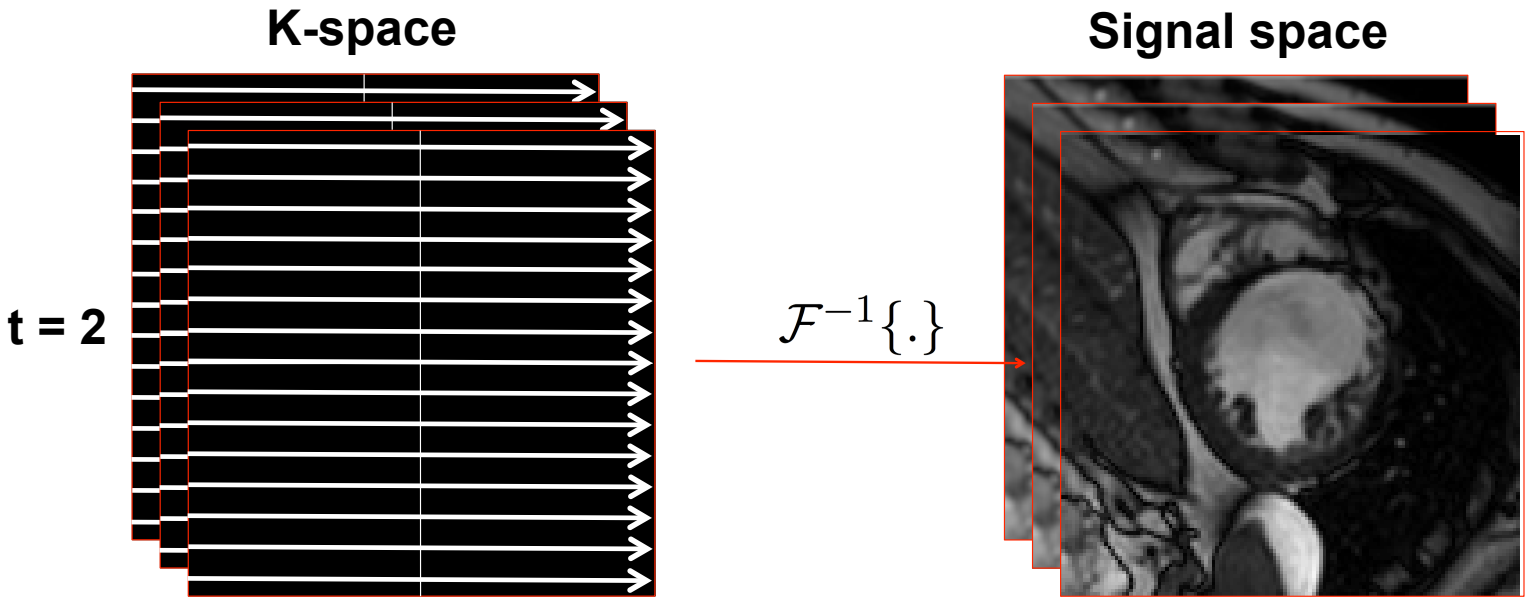
Cardiac MRI: Full acquisition is slow

- MRI acquisition is performed in k-space by sequentially traversing sampling trajectories.



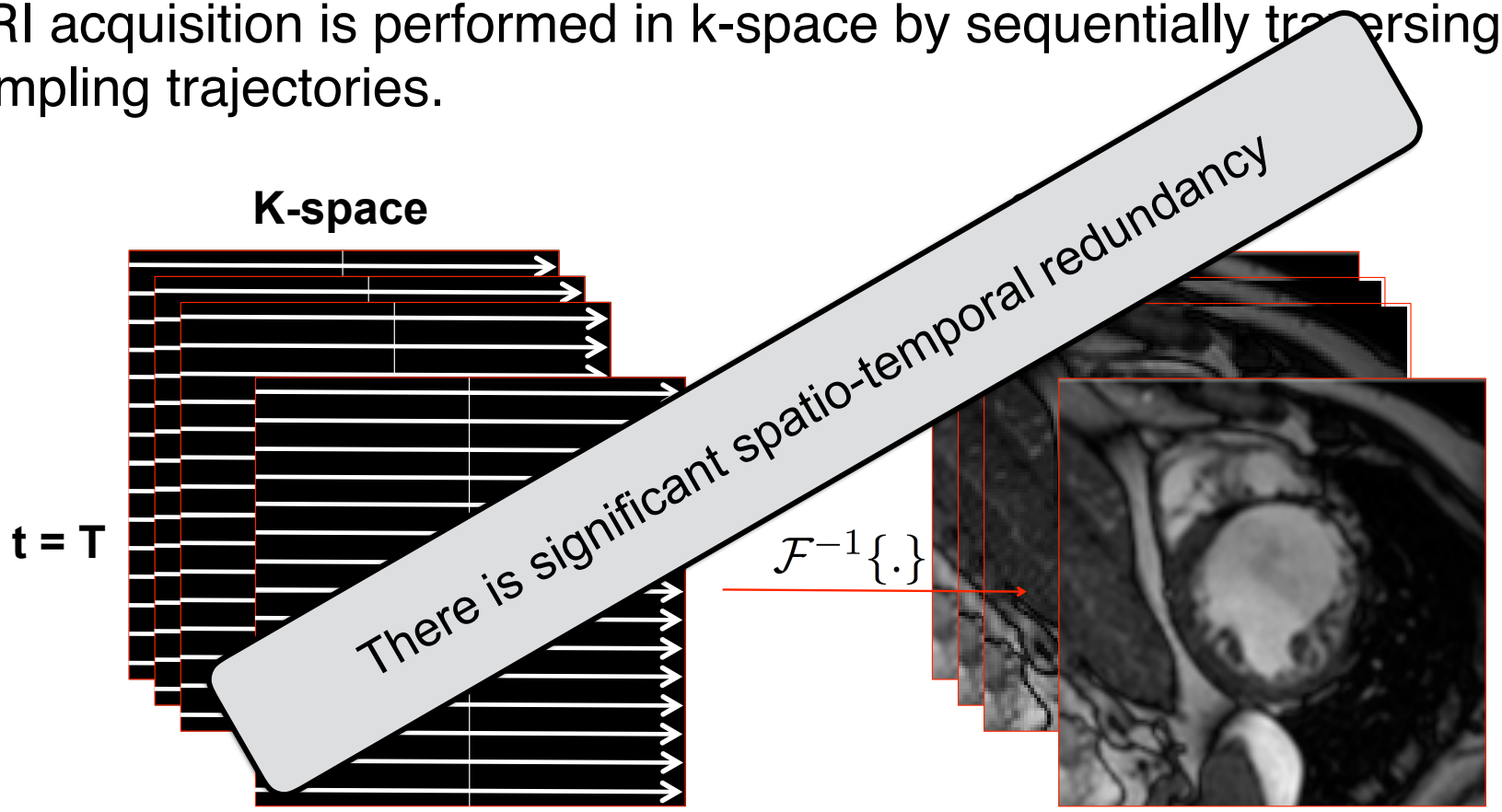
Cardiac MRI: Full acquisition is slow

- MRI acquisition is performed in k-space by sequentially traversing sampling trajectories.



Cardiac MRI: Full acquisition is slow

- MRI acquisition is performed in k-space by sequentially traversing sampling trajectories.



K-space undersampling

- Acquiring a fraction of k-space **accelerates** the process but introduces **aliasing** in signal space.

K-space undersampling

- Acquiring a fraction of k-space **accelerates** the process but introduces **aliasing** in signal space.

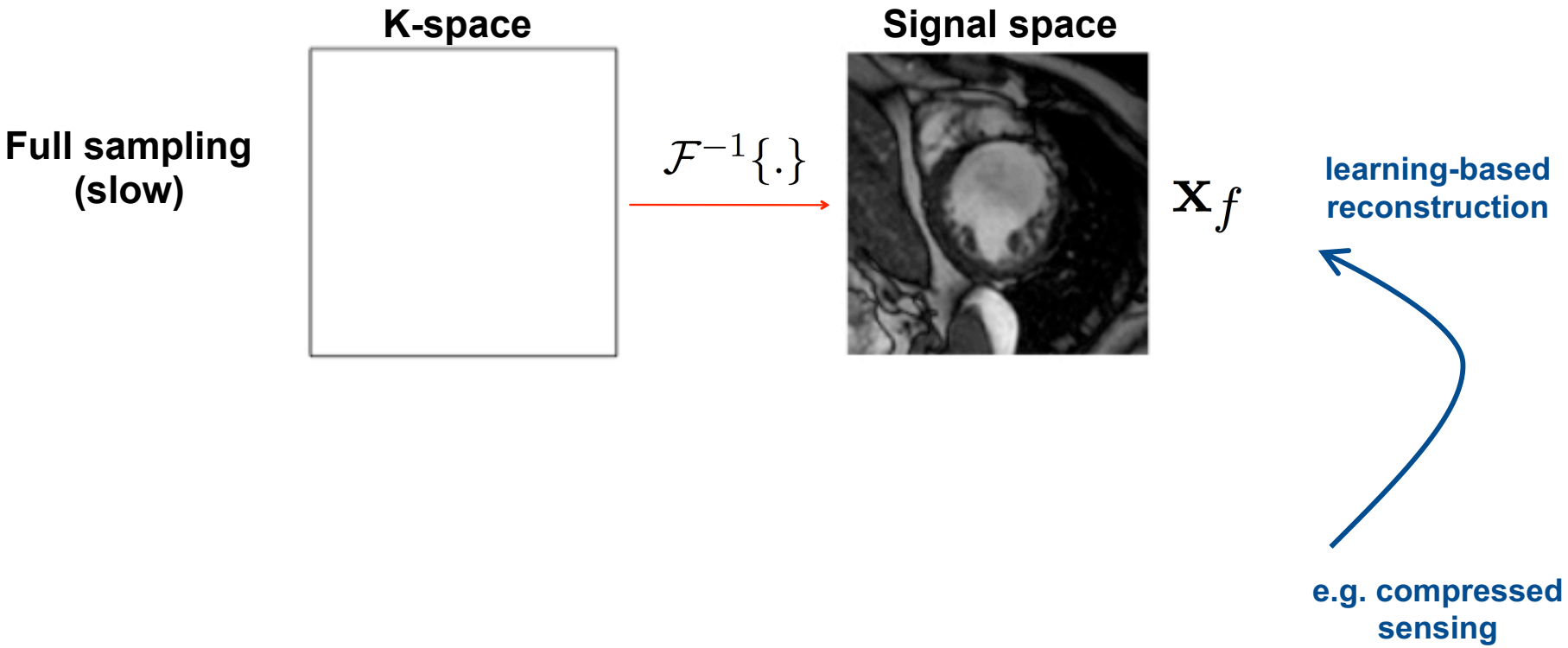


Image reconstruction from undersampled k-space

- One can recover full k-space through compressed sensing techniques:
  - Lustig et al., MRM 2007
  - Jung et al., MRM 2009
  - Otazo et al., MRM 2010

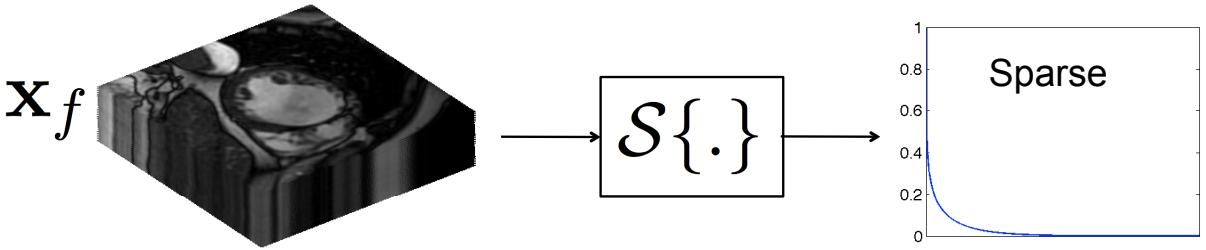
Based on generic priors, e.g. sparsity or low-rank

- More recently other techniques have shown to be powerful for this task as well:
  - Caballero et al., IEEE TMI 2014: Dictionary learning
  - Bhatia et al., MICCAI 2016: Manifold learning
  - Schlemper et al., IEEE TMI 2017: Deep learning for cardiac MRI
  - Hammernik et al., MRM 2017: Deep learning for knee MRI
  - Ye et al., SIAM IS 2018: Deep convolutional framelets

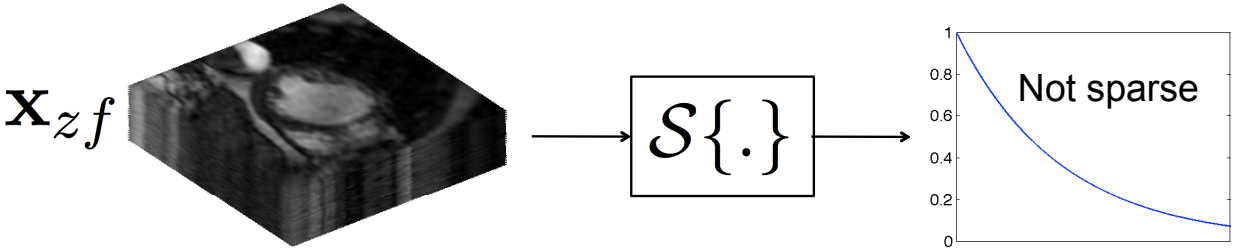
Based on learnt priors

Sparsity

- Most natural signals are compressible under some domain.



- Aliasing makes this assumption break down, so it can be imposed on the reconstruction of a signal.



Compressed sensing

- Assume $\hat{\mathbf{x}}_u$ is the undersampled observation in k-space and $\mathcal{F}_u$ is the undersampled Fourier operator.
- We look for solution $\mathbf{x}$ such that:

Compressed sensing

- Assume $\hat{\mathbf{x}}_u$ is the undersampled observation in k-space and $\mathcal{F}_u$ is the undersampled Fourier operator.
- We look for solution $\mathbf{x}$ such that:
  - It is consistent with k-space observation

$$\|\mathcal{F}_u\{\mathbf{x}\} - \hat{\mathbf{x}}_u\|_2^2 < \epsilon$$

Compressed sensing

- Assume $\hat{\mathbf{x}}_u$ is the undersampled observation in k-space and $\mathcal{F}_u$ is the undersampled Fourier operator.
- We look for solution $\mathbf{x}$ such that:
  - It is consistent with k-space observation $\mathbf{x}$
  - It has the sparsest representation under $\mathcal{S}\{\mathbf{x}\}$.

$$\min_{\mathbf{x}} \|\mathcal{S}\{\mathbf{x}\}\|_0$$

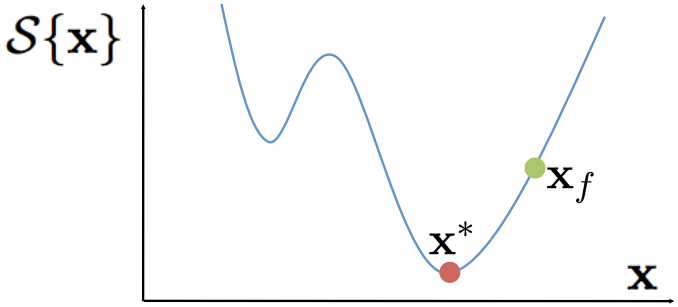
$$\|\mathcal{F}_u\{\mathbf{x}\} - \hat{\mathbf{x}}_u\|_2^2 < \epsilon$$

Compressed sensing

- Assume $\hat{\mathbf{x}}_u$ is the undersampled observation in k-space and $\mathcal{F}_u$ is the undersampled Fourier operator.
- We look for solution $\mathbf{x}$ such that:
  - It is consistent with k-space observation $\mathbf{x}$
  - It has the sparsest representation under $\mathcal{S}\{\mathbf{x}\}$.

Optimisation problem:

$$\min_{\mathbf{x}} \|\mathcal{S}\{\mathbf{x}\}\|_0 \qquad s.t. \qquad \|\mathcal{F}_u\{\mathbf{x}\} - \hat{\mathbf{x}}_u\|_2^2 < \epsilon$$



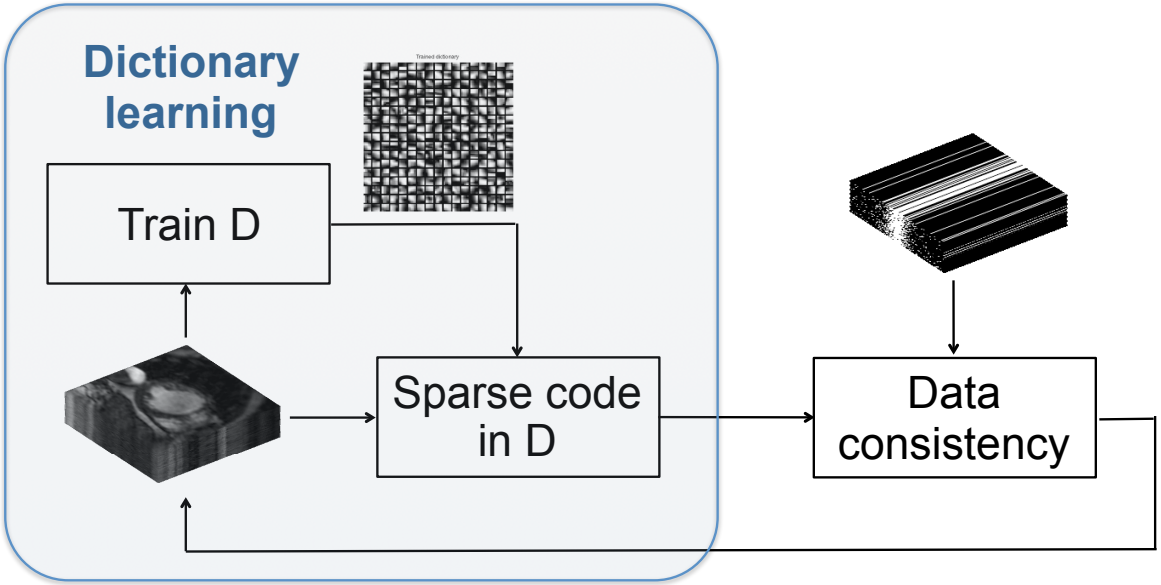
Need to ensure that transform $\mathcal{S}\{\mathbf{x}\}$ really does provide the sparsest representation in the solution space generated.

Compressed sensing and dictionary learning

- Many compressed sensing approaches used fixed sparsity transforms (e.g. wavelets) for $\mathcal{S}\{\mathbf{x}\}$
- Adaptive sparsity transforms (or dictionary learning) can learn optimal sparsity transforms from the data
- Has been successfully used for CS-MRI
  - Ravishankar & Bresler, IEEE TMI 2011
  - Awate et al., IEEE ISBI 2012
  - Doneva et al., MRM 2010
- Can be extended to dynamic MRI

Dictionary learning for MR reconstruction

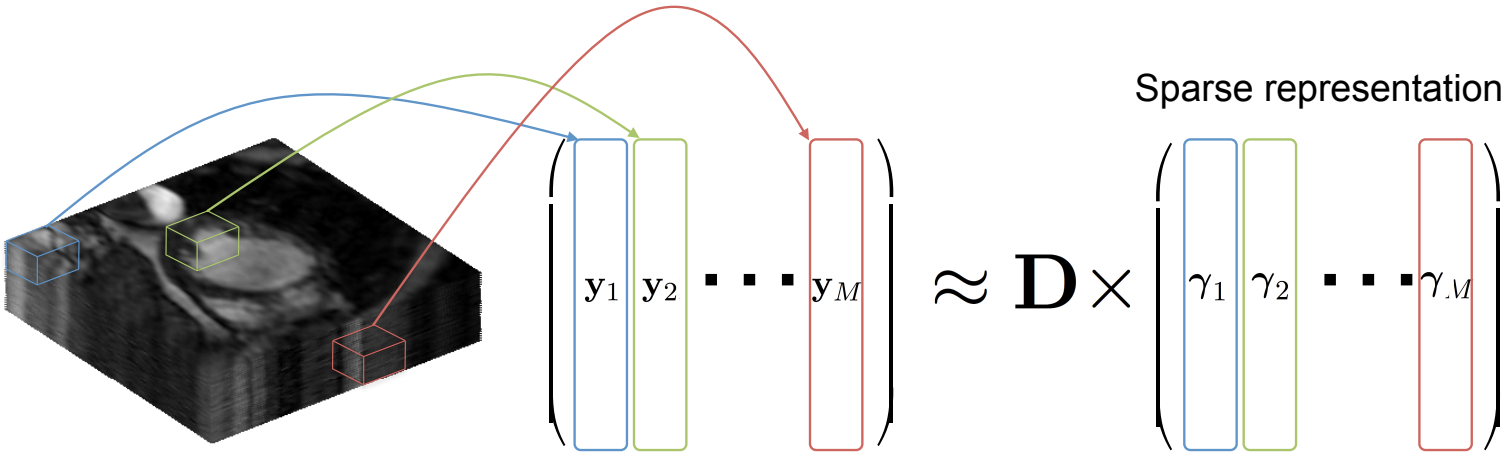
- Objective: Out of all solutions consistent with the acquired k-space, we look for the one that is sparsest under the learned dictionary.



Caballero et al., IEEE TMI 2014

Step 1: Dictionary learning

- **<u>Training:</u>** Learn a dictionary that will sparsely represent 3D patches randomly extracted from the corrupted sequence.

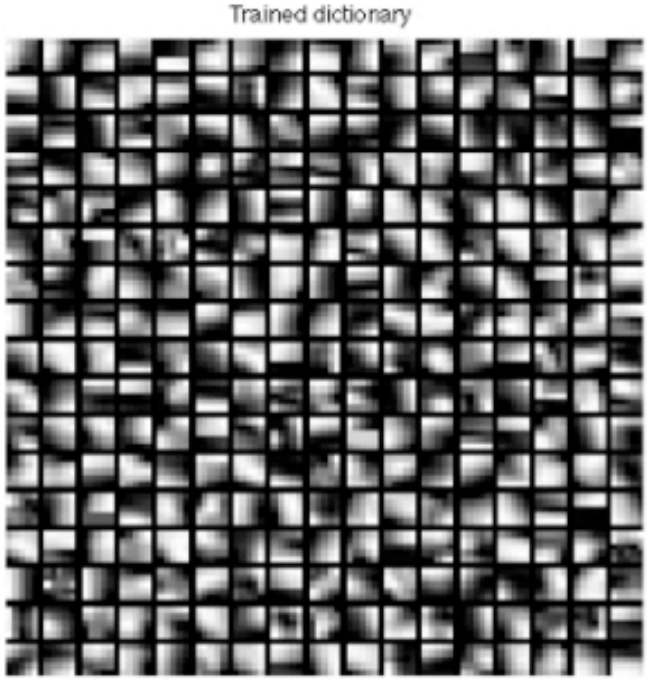


$$\min_{\mathbf{\Gamma},\mathbf{D}} \|\boldsymbol{\gamma}_i\|_0 \qquad s.t. \qquad \|\mathbf{y}_i - \mathbf{D}\boldsymbol{\gamma}_i\|_2^2 < \epsilon, \forall i$$

Caballero et al., IEEE TMI 2014

Dictionary learning – Example

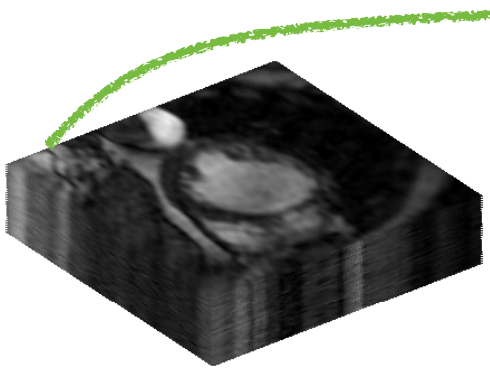
- The dictionary is adapted to features in the data and by construction provides a sparse representation of it.



Caballero et al., IEEE TMI 2014

Step 2: Sparse coding

- **Coding:** The entire sequence is sparsely coded using $\mathbf{D}$.



$$\min_{\boldsymbol{\Gamma}} \sum_{j=1}^{P} \|\boldsymbol{\gamma}_j\|_0 \quad s.t. \quad \sum_{j=1}^{P} \|\mathbf{R}_j\mathbf{y} - \mathbf{D}\boldsymbol{\gamma}_j\|_2^2 < \epsilon_1$$

- The sparse coding $\boldsymbol{\Gamma}$ provides an approximation of the sequence excluding part of the aliasing.

$\mathbf{D}\boldsymbol{\Gamma}$

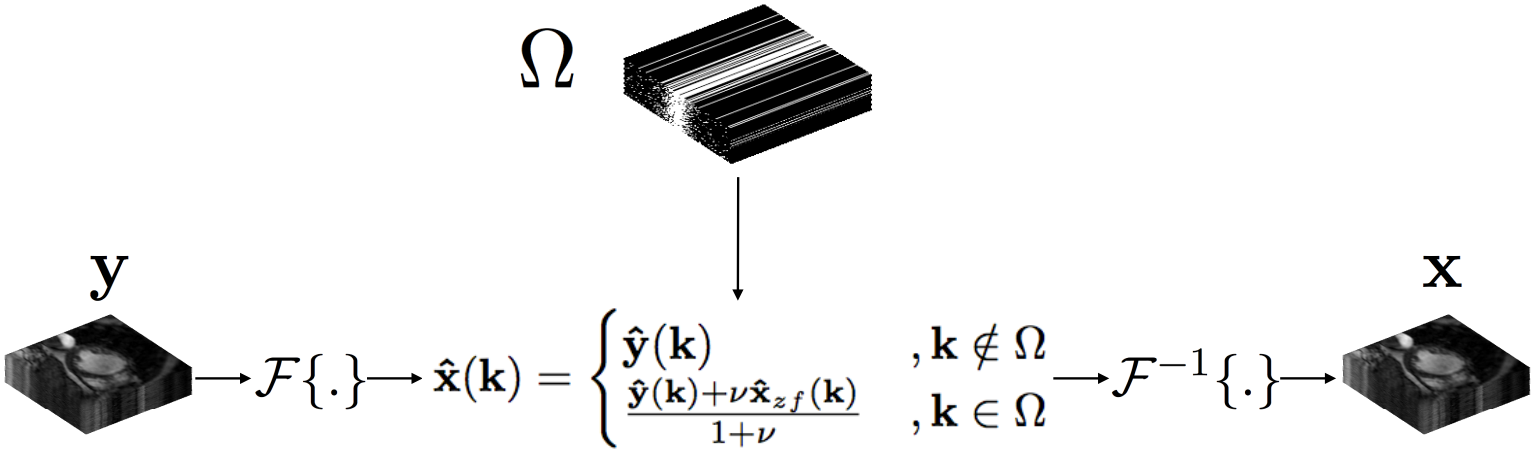
$\boldsymbol{\Gamma}$: Sparse representation under $\mathbf{D}$

$\mathbf{D}\boldsymbol{\Gamma}$: Sparse approximation of $\mathbf{y}$

Caballero et al., IEEE TMI 2014

Step 3: Data consistency

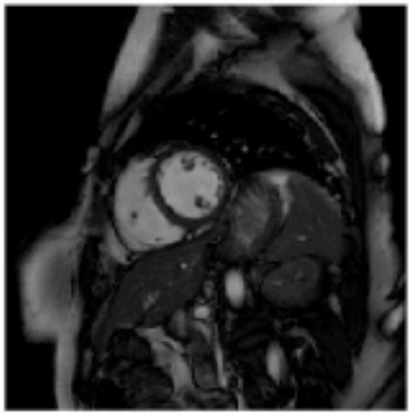
- Processing in signal space will make the k-space of solution different from the initial observations.
- Data consistency in k-space must be enforced.



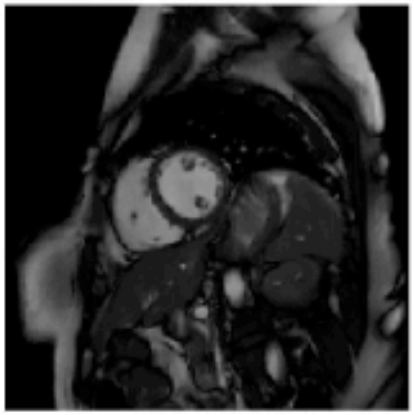
Caballero et al., IEEE TMI 2014

Magnitude reconstruction (8-fold)

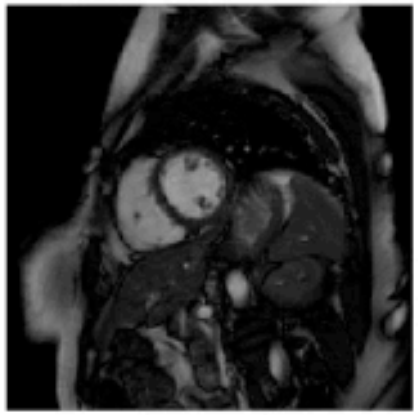
-]



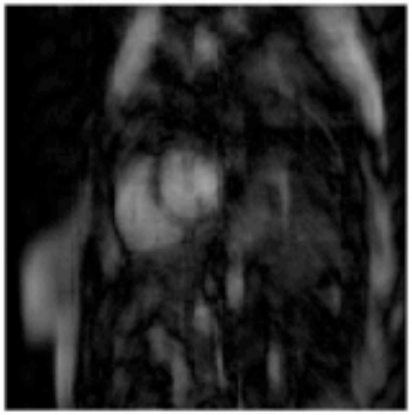
Original



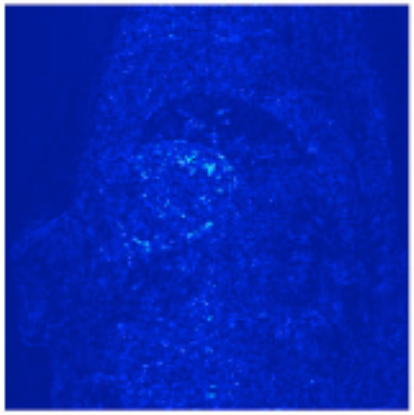
DLTG (36.7 dB)



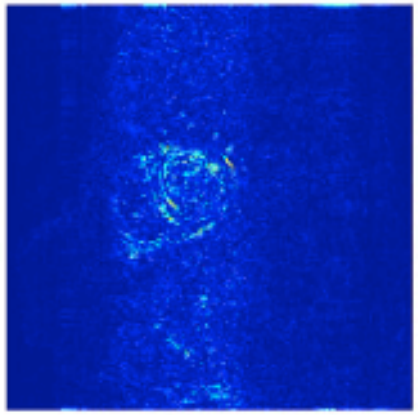
k-t FOCUSS (34.3 dB)



Zero-filled (22.7 dB)



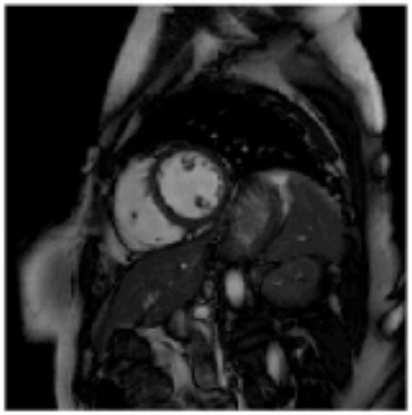
DLTG error x5



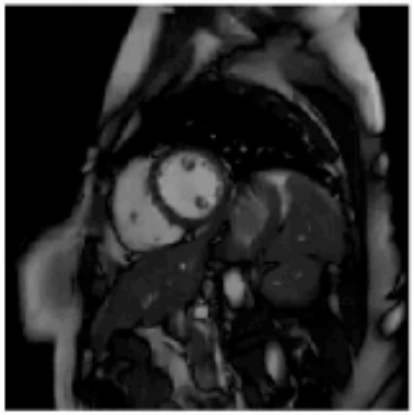
k-t FOCUSS error x5

Caballero et al., IEEE TMI 2014

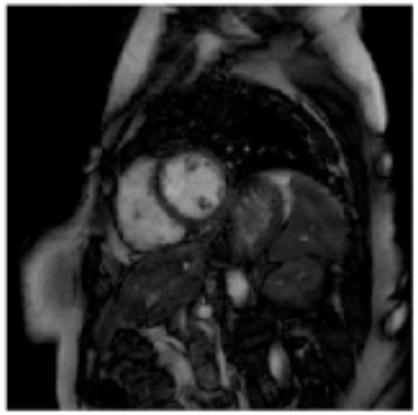
Magnitude reconstruction (12-fold)



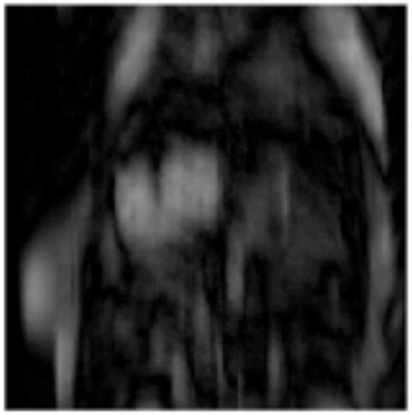
Original



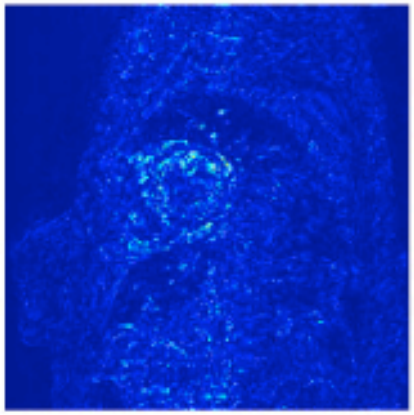
DLTG (34.0 dB)



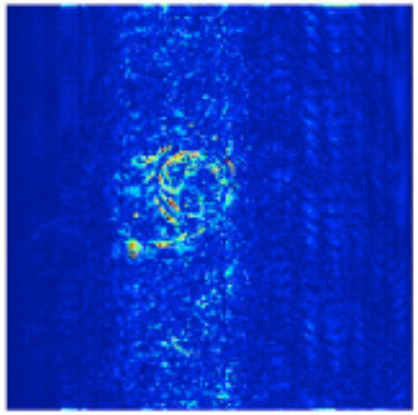
k-t FOCUSS (31.4 dB)



Zero-filled (21.9 dB)



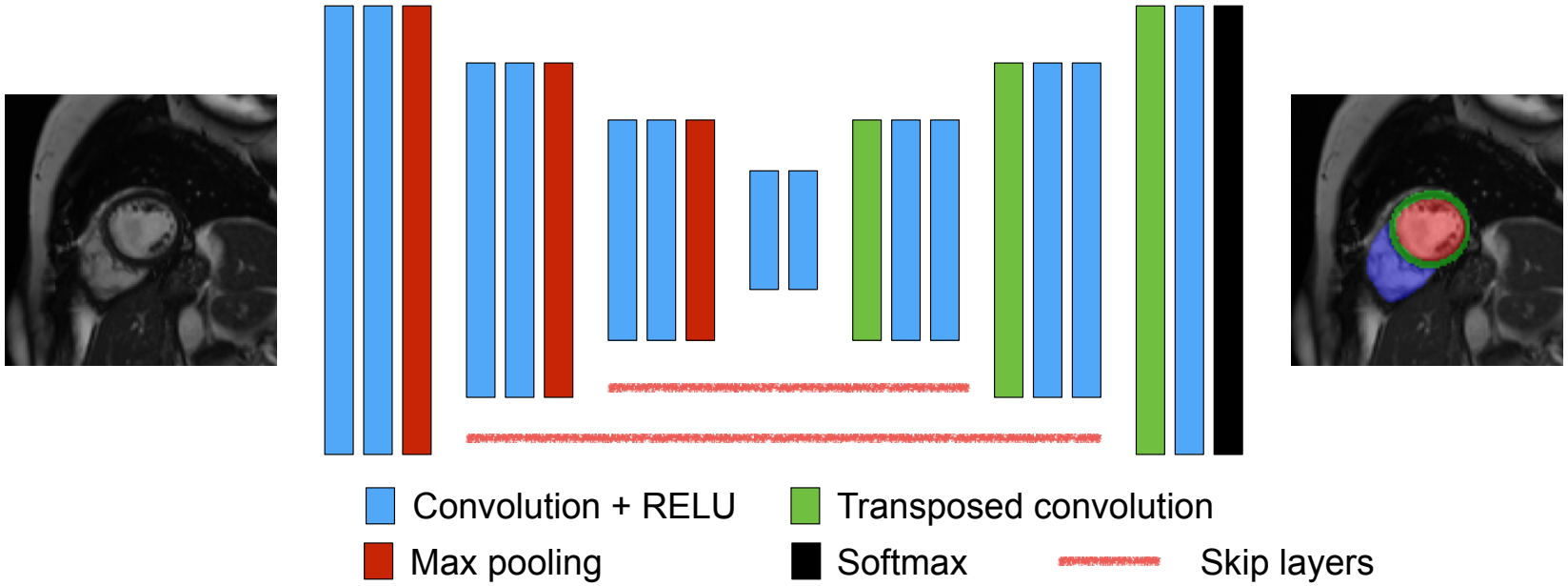
DLTG error x5



k-t FOCUSS error x5

Caballero et al., IEEE TMI 2014

What about deep learning?



Problem formulation

- Reconstruct image $\mathbf{x} \in \mathbb{C}^N$ given undersampled k-space measurements $\mathbf{y} \in \mathbb{C}^M$ $(M \ll N)$:

$$\mathbf{y} = \mathbf{F}_u \mathbf{x} + \mathbf{e}$$

Undersampled Fourier encoding matrix

Acquisition noise

Problem formulation

- Reconstruct image $\mathbf{x} \in \mathbb{C}^N$ given undersampled k-space measurements $\mathbf{y} \in \mathbb{C}^M$ $(M \ll N)$:

$$\mathbf{y} = \mathbf{F}_u \mathbf{x} + \mathbf{e}$$

Undersampled Fourier encoding matrix

Acquisition noise

- In the case of Cartesian sampling we have $\mathbf{F}_u = \mathbf{M}\mathbf{F}$ where $\mathbf{F} \in \mathbb{C}^{N \times N}$ applies the 2D Fourier transform and $\mathbf{M} \in \mathbb{C}^{M \times N}$ is the undersampling mask in k-space

Problem formulation

- We are trying to solve the following unconstrained optimisation problem:

$$\min_{\mathbf{x}} \mathcal{R}(\mathbf{x}) + \lambda \|\mathbf{y} - \mathbf{F}_u \mathbf{x}\|_2^2$$

Regularisation term
(in CS usually the l_0 or l_1 norm)

Data fidelity term

Problem formulation

- We are trying to solve the following unconstrained optimisation problem:

$$\min_{\mathbf{x}} \mathcal{R}(\mathbf{x}) + \lambda \|\mathbf{y} - \mathbf{F}_u \mathbf{x}\|_2^2$$

Regularisation term
(in CS usually the l_0 or l_1 norm)

Data fidelity term

- For CNN based reconstruction we formulate the problem as

$$\min_{\mathbf{x}} \|\mathbf{x} - f_{\text{cnn}}(\mathbf{x}_u|\boldsymbol{\theta})\|_2^2 + \lambda \|\mathbf{F}_u \mathbf{x} - \mathbf{y}\|_2^2$$

Data consistency layer

- To ensure data fidelity, we add a data consistency layer. For fixed network parameters we can write:

$$\mathbf{s}_{\text{rec}}(j) = \begin{cases} \mathbf{s}_{\text{cnn}}(j) & \text{if } j \notin \Omega \\ \dfrac{\mathbf{s}_{\text{cnn}}(j) + \lambda \mathbf{s}_0(j)}{1+\lambda} & \text{if } j \in \Omega \end{cases}$$

Missing part of k-space

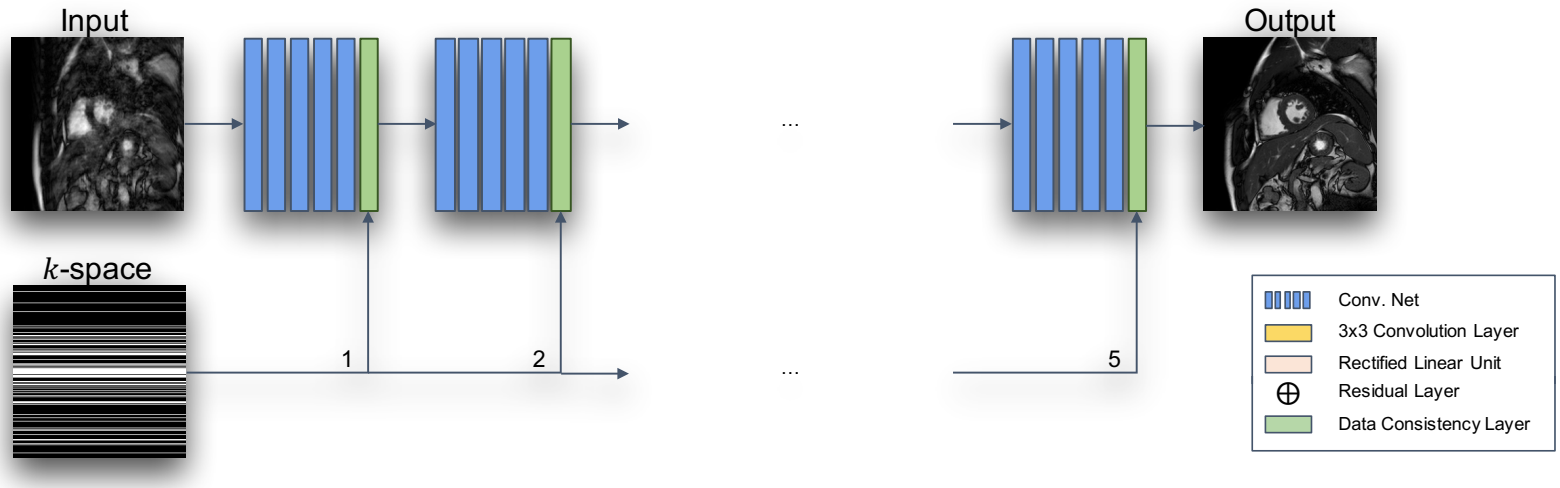
Acquired part of k-space

Fourier-encoding of reconstructed image

Zero-filled k-space

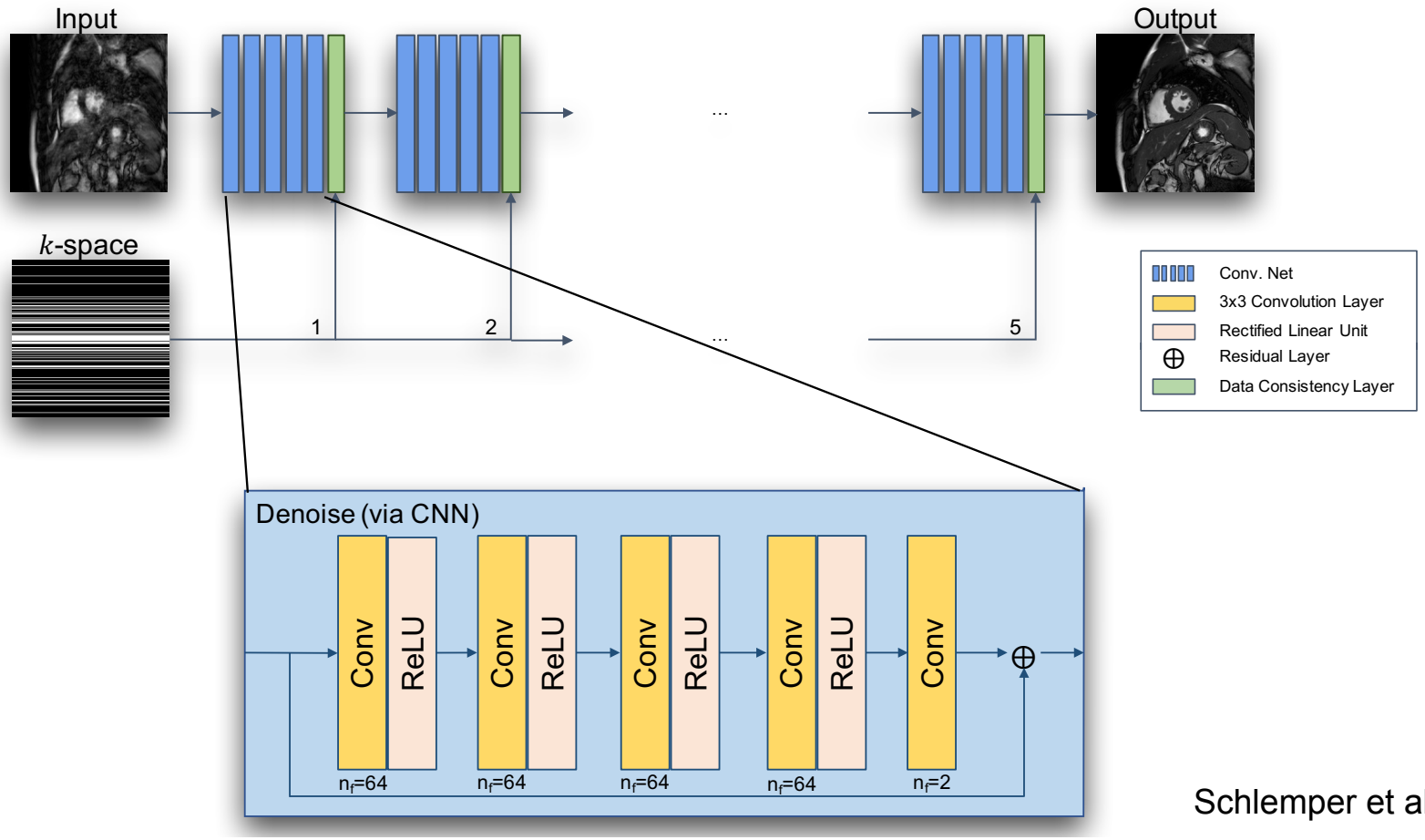
$$\mathbf{s}_{\text{cnn}} = \mathbf{F}\mathbf{x}_{\text{cnn}} = \mathbf{F} f_{\text{cnn}}(\mathbf{x}_u|\boldsymbol{\theta})$$

Deep Cascade of CNNs for MRI Reconstruction



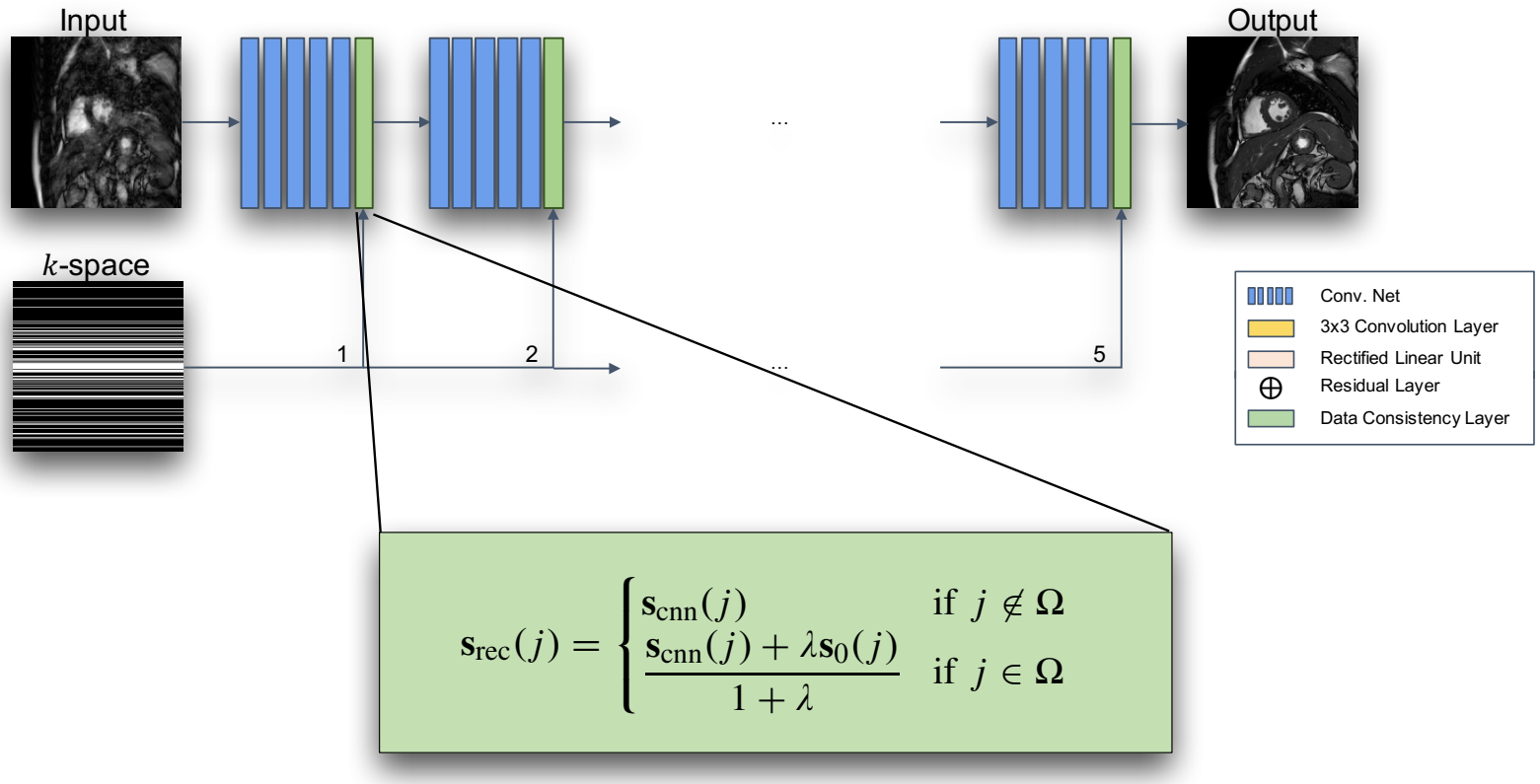
Schlemper et al. IEEE TMI 2017

Deep Cascade of CNNs for MRI Reconstruction



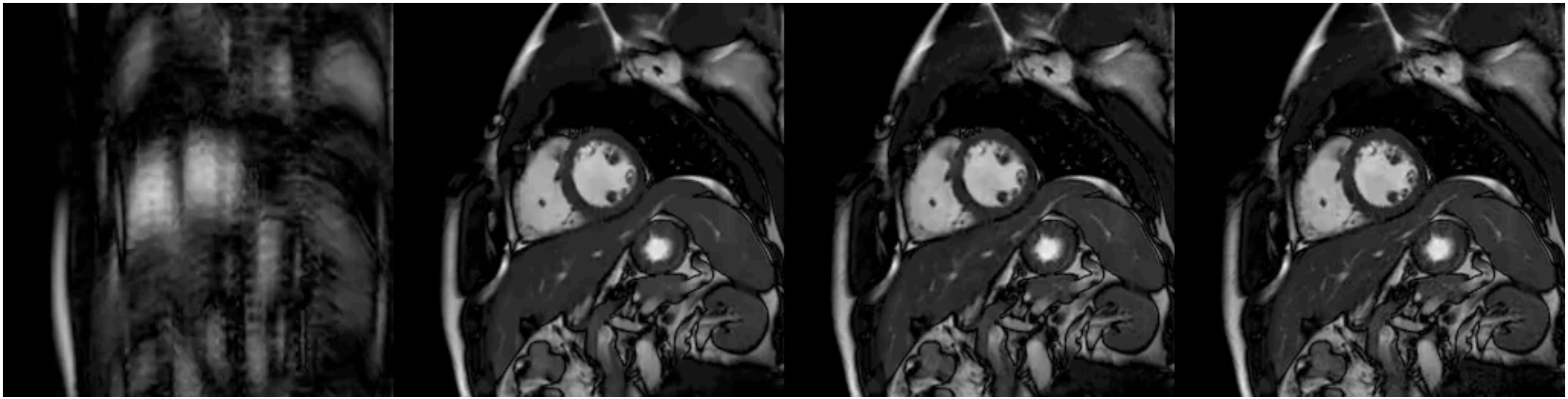
Schlemper et al. IEEE TMI 2017

Deep Cascade of CNNs for MRI Reconstruction



Schlemper et al. IEEE TMI 2017

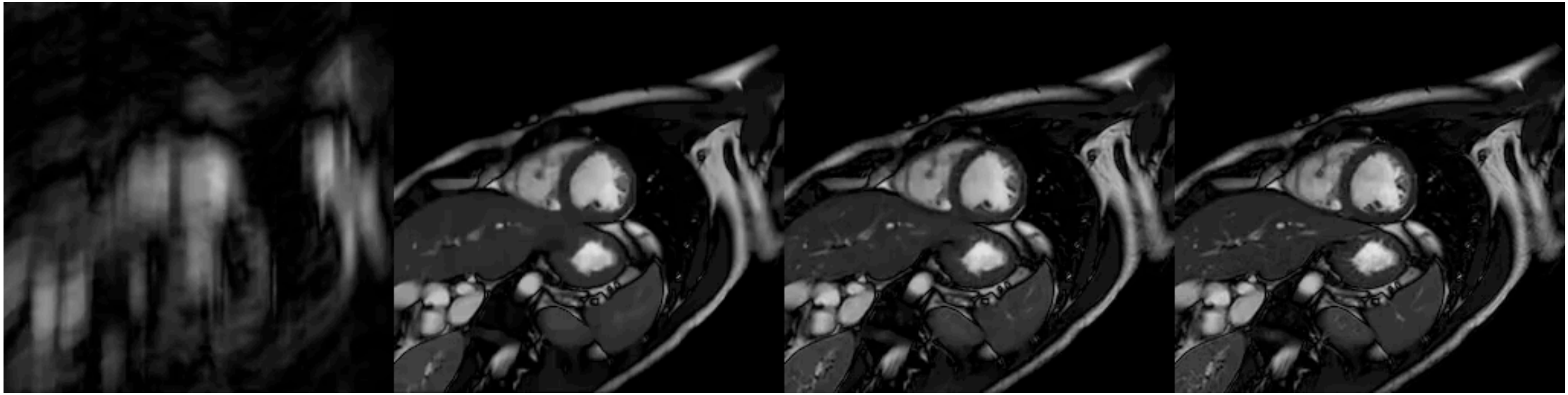
Magnitude reconstruction (6-fold)



(a) 6x Undersampled **(b) DLTG** **(c) CNN** **(d) Ground Truth**

Schlemper et al. IEEE TMI 2017

Magnitude reconstruction (11-fold)



(a) 11x Undersampled **(b) DLTG** **(c) CNN** **(d) Ground Truth**

Schlemper et al. IEEE TMI 2017

Deep Cascade of CNNs for MRI Reconstruction: Results

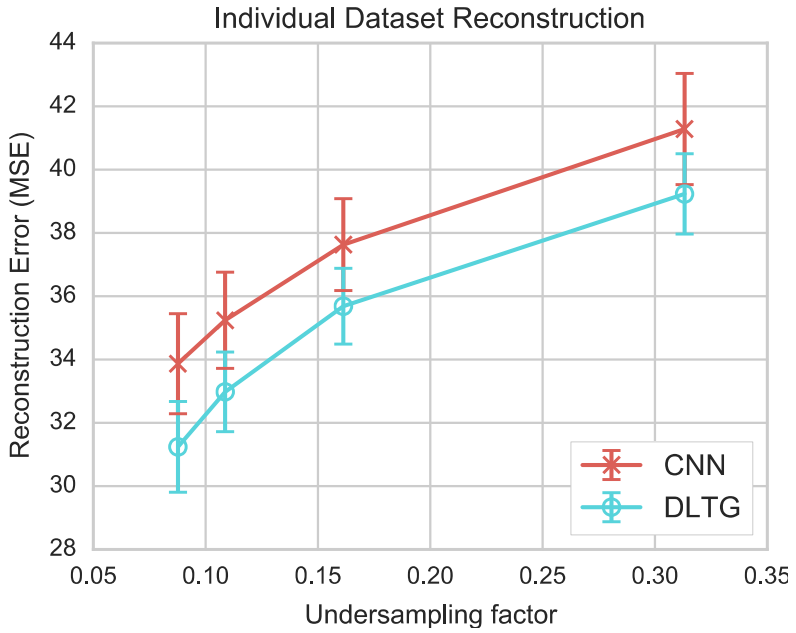
- Test error across 10 subjects:

PSNR

Model	R=4 (dB)	R=8 (dB)
DLTG	27.5 (1.31)	22.6 (0.95)
CNN	**31.0 (1.08)**	**25.2 (1.00)**

Speed

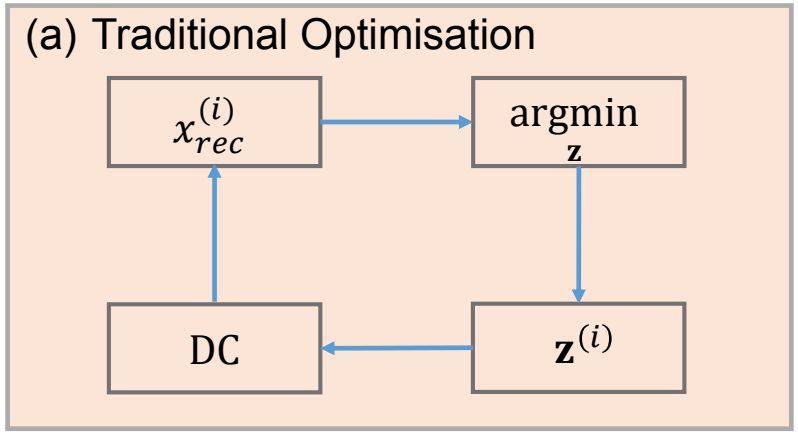
Model	Time
DLMRI/DLTG	~6 hr (CPU)
CNN (2D)	0.69 s (GPU)
CNN (2D+t)	10 s (GPU)



2D+t (vs. DLTG)

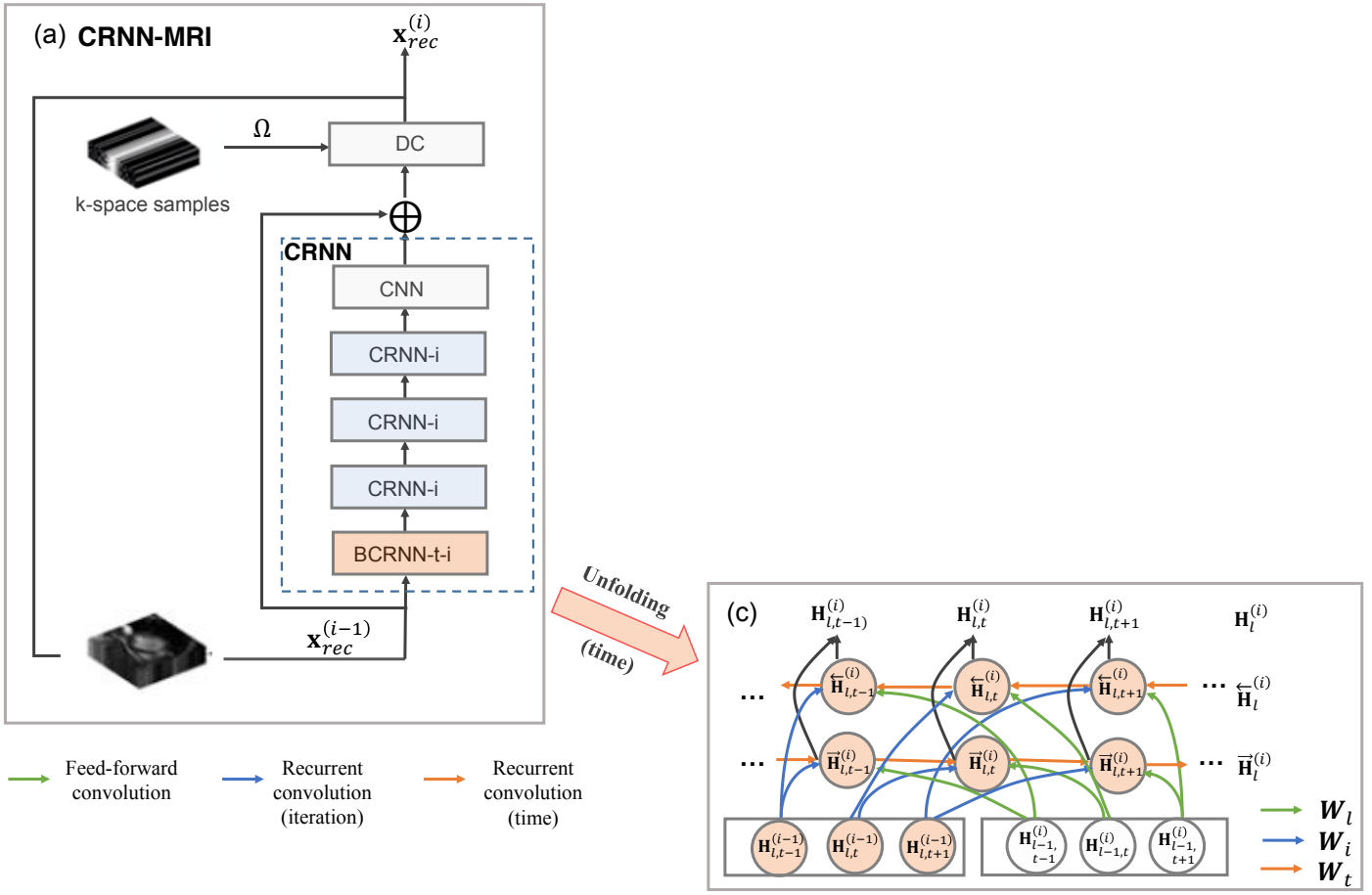
Schlemper et al. IEEE TMI 2017

Replacing Deep Cascade of CNNs with Recurrent Neural Networks



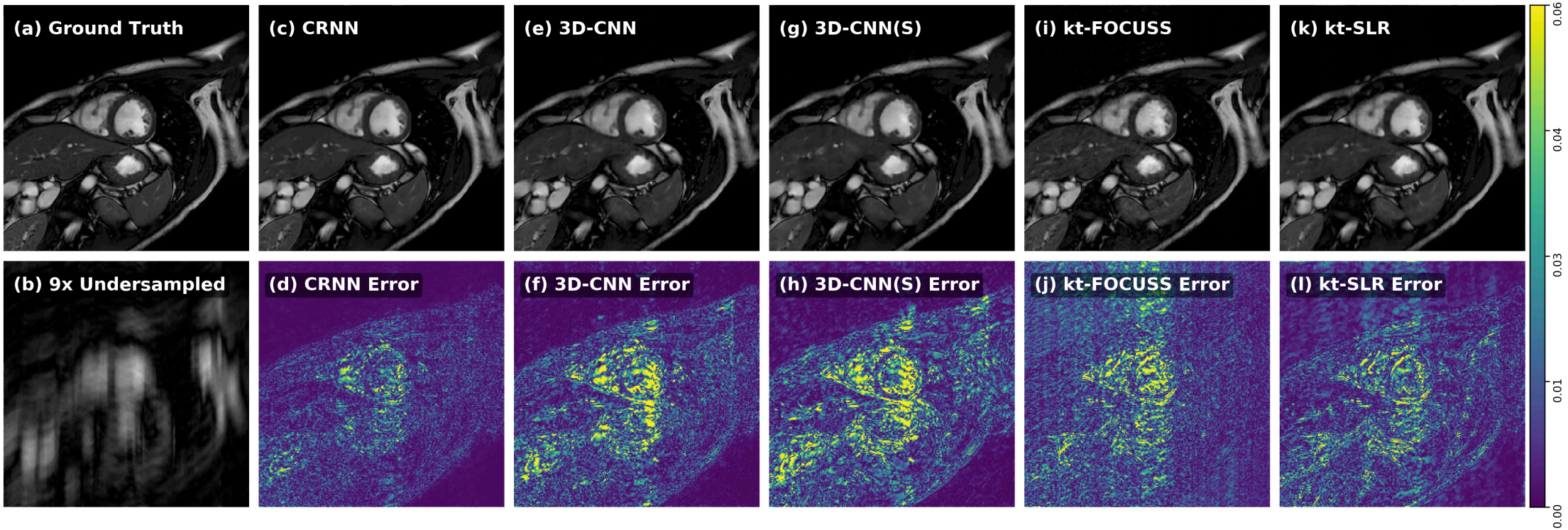
C. Qin et al. arXiv:1712.01751, in press, IEEE TMI 2018

Replacing Deep Cascade of CNNs with Recurrent Neural Networks



C. Qin et al. arXiv:1712.01751, in press, IEEE TMI 2018

Replacing Deep Cascade of CNNs with Recurrent Neural Networks



C. Qin et al. arXiv:1712.01751, in press, IEEE TMI 2018

Overview

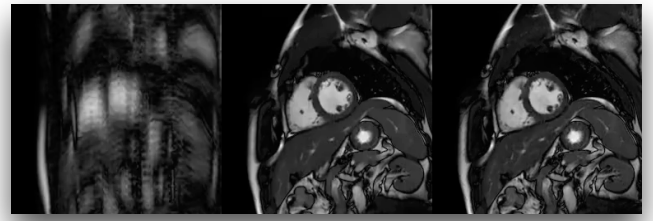


Image reconstruction

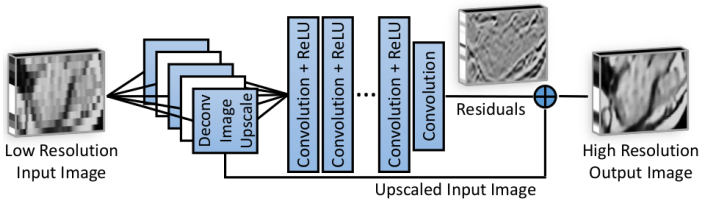


Image super-resolution

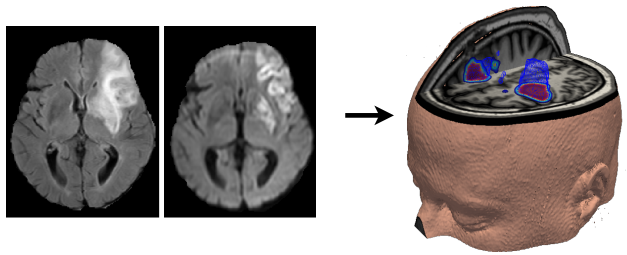


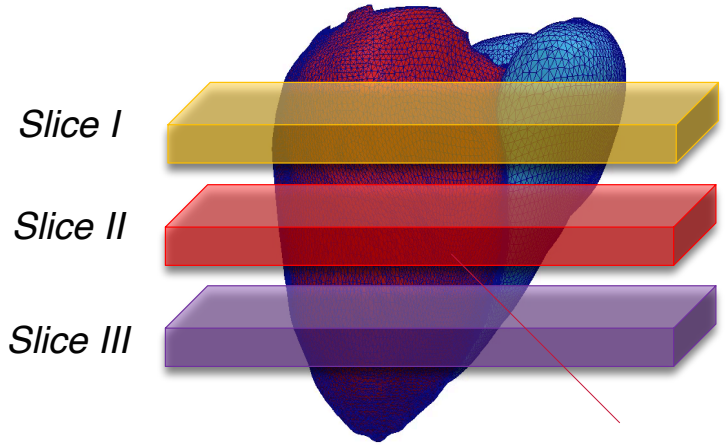
Image segmentation

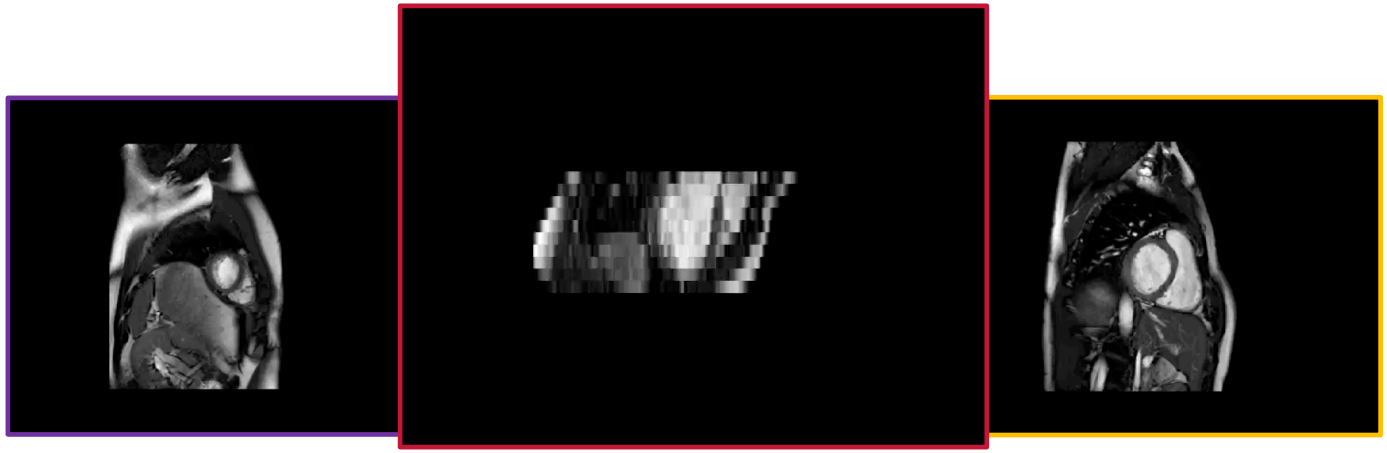
Abdominal View
Confidence: 98%

Image classification

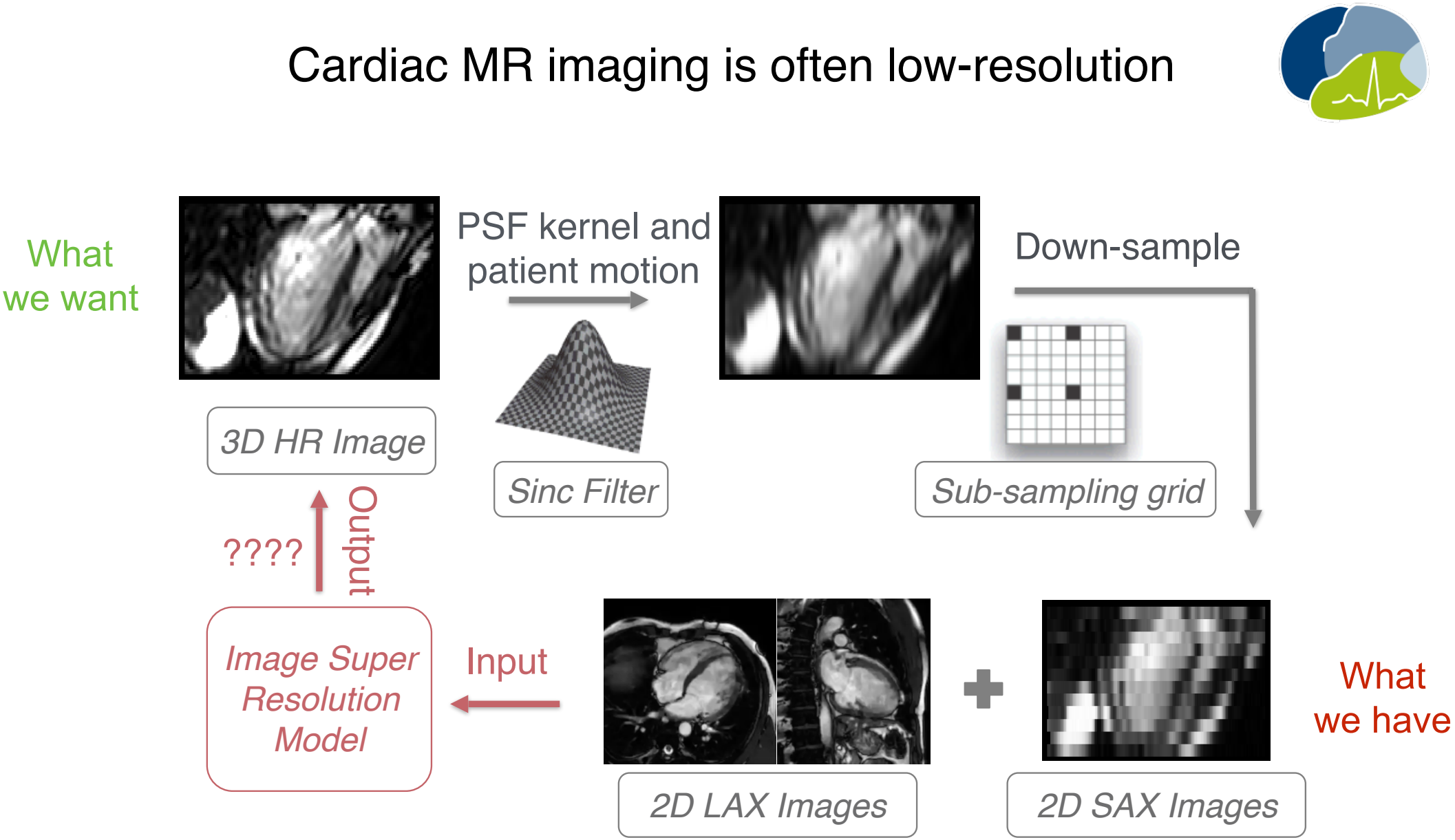
Cardiac MR imaging is often low-resolution

- Acquisition of cardiac MRI typically consists of 2D multi-slice data due to
  - constraints on SNR
  - breath-hold time
  - total acquisition time
- This leads to thick slice data (thickness 8-10 mm per slice)





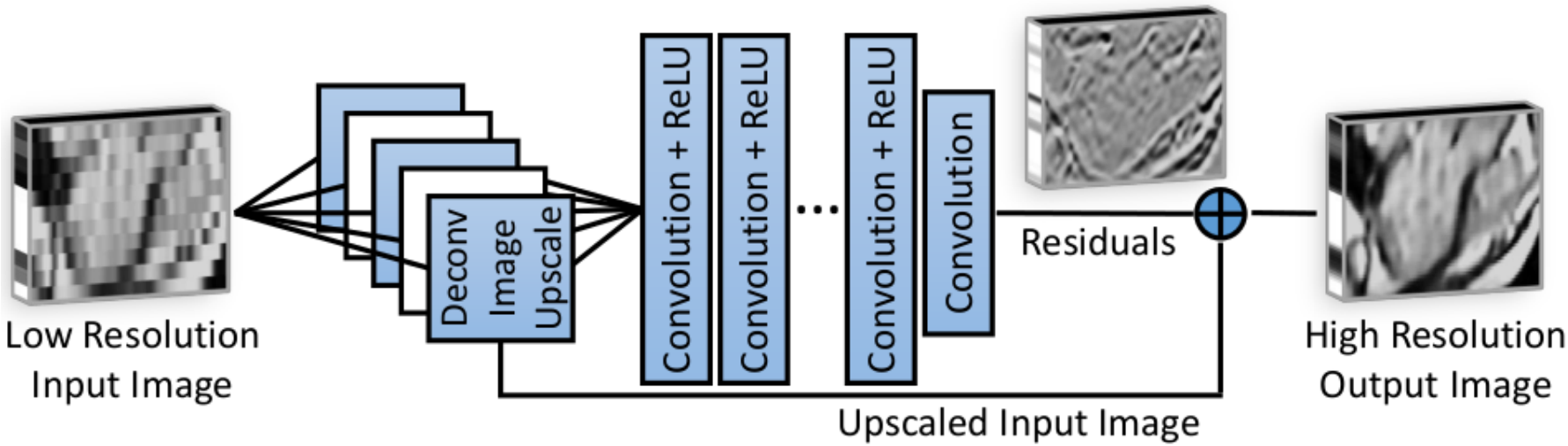
Cardiac MR imaging is often low-resolution



Super-resolution for cardiac MR imaging

Components of the model

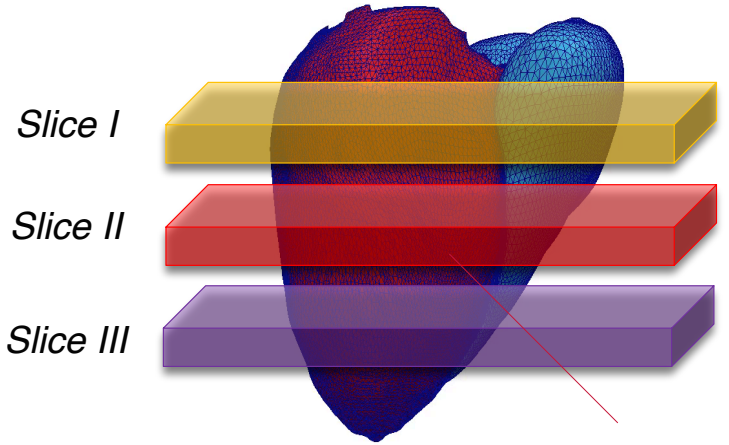
- 3D Convolution and Deconvolution (inverse convolution) Kernels
-
-
-

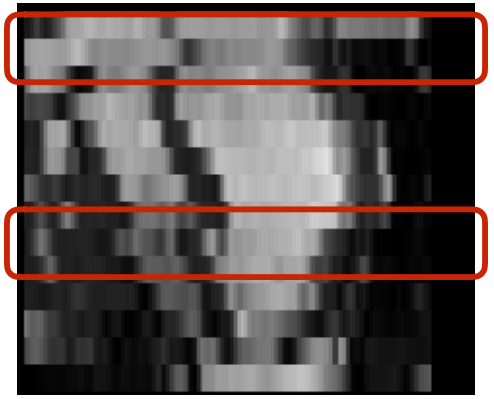


O. Oktay et al. MICCAI 2016

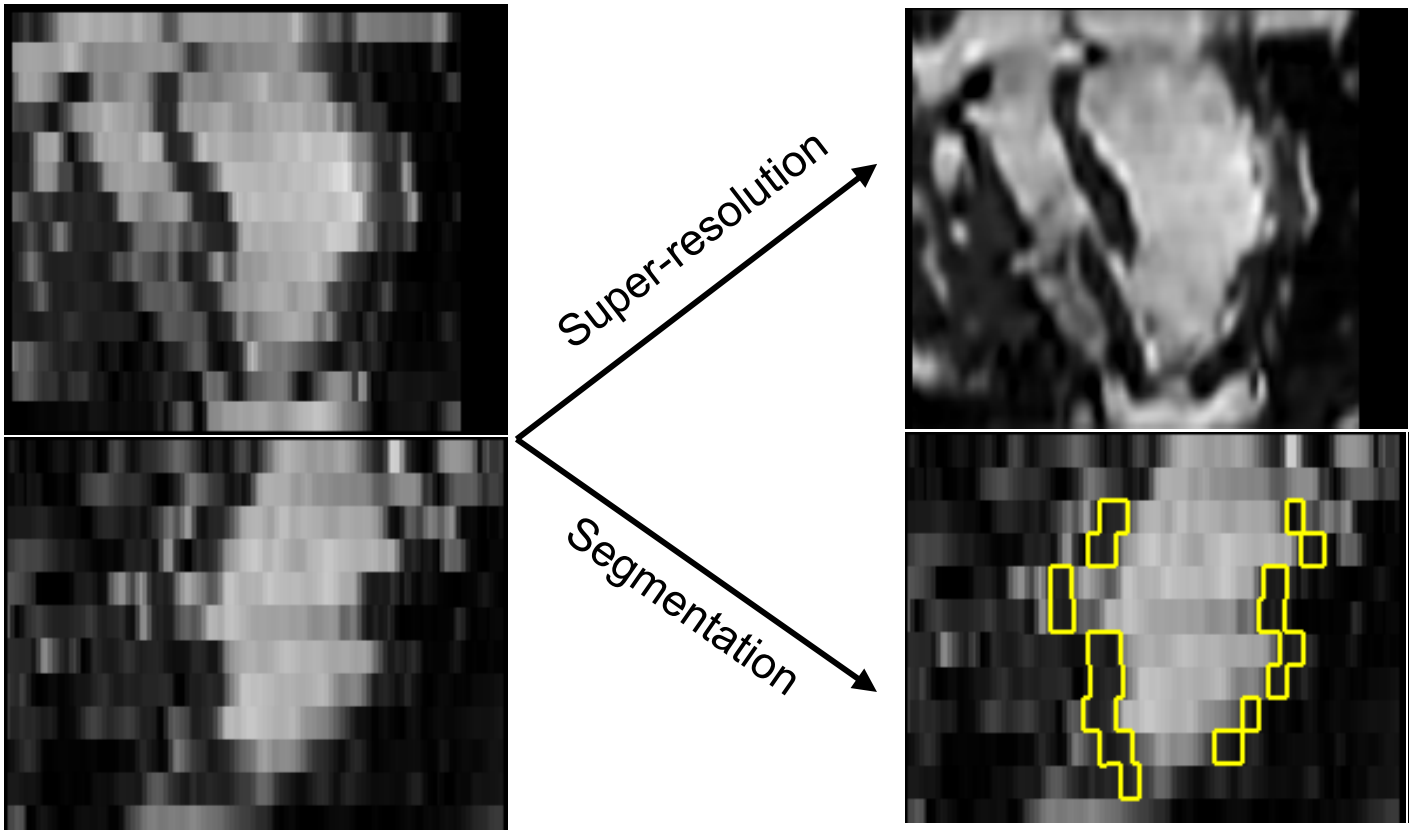
But: Cardiac MR imaging is still challenging

- Acquisition of cardiac MRI typically consists of 2D multi-slice data due to
  - constraints on SNR
  - breath-hold time
  - total acquisition time
- This leads to thick slice data (thickness 8-10 mm per slice)
- Motion between slices can lead to artefacts

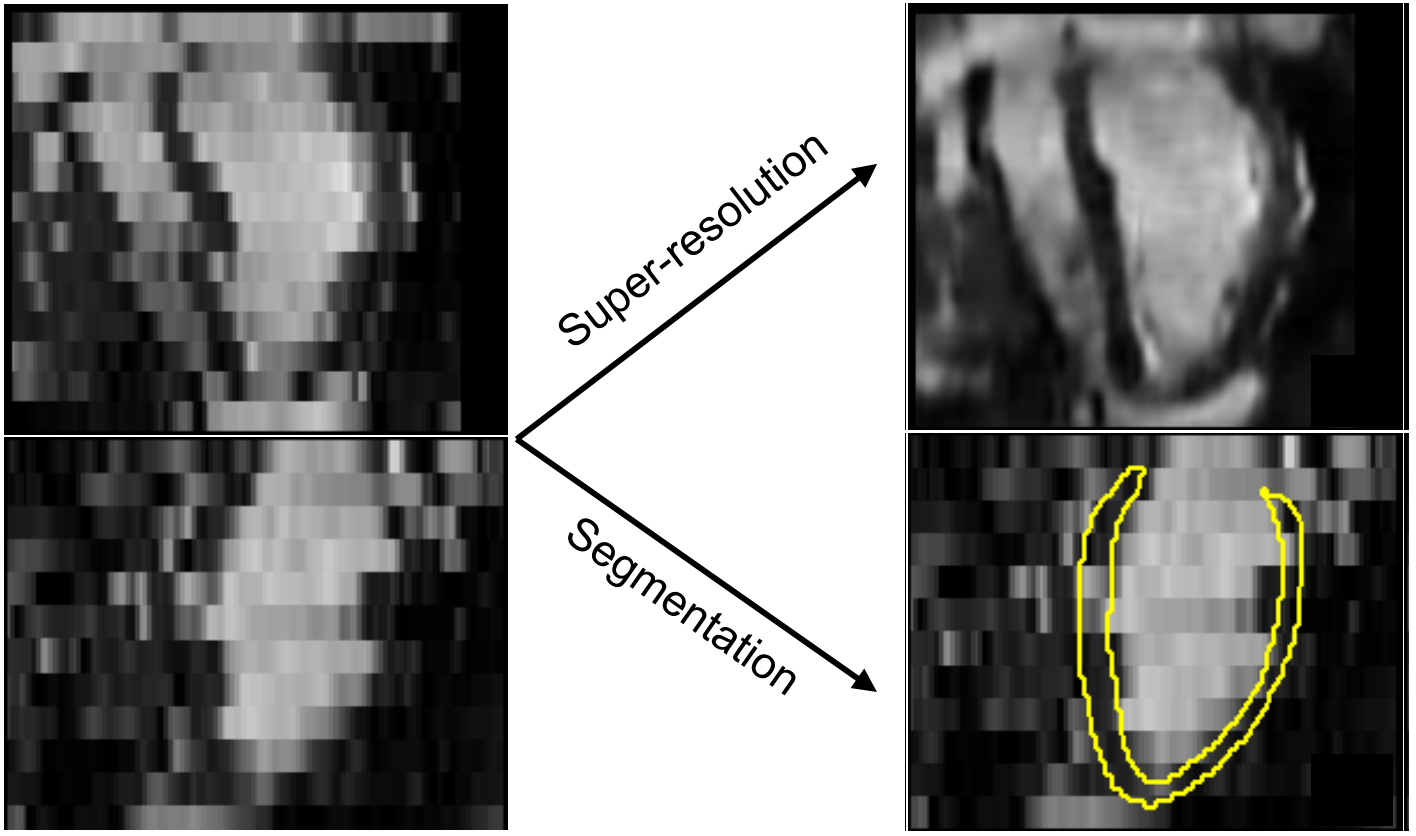




Conventional CNNs: Problem



Conventional CNNs: What we want



Conventional CNNs: No explicit use of prior knowledge

- Standard Loss for **segmentation**: Cross-Entropy loss

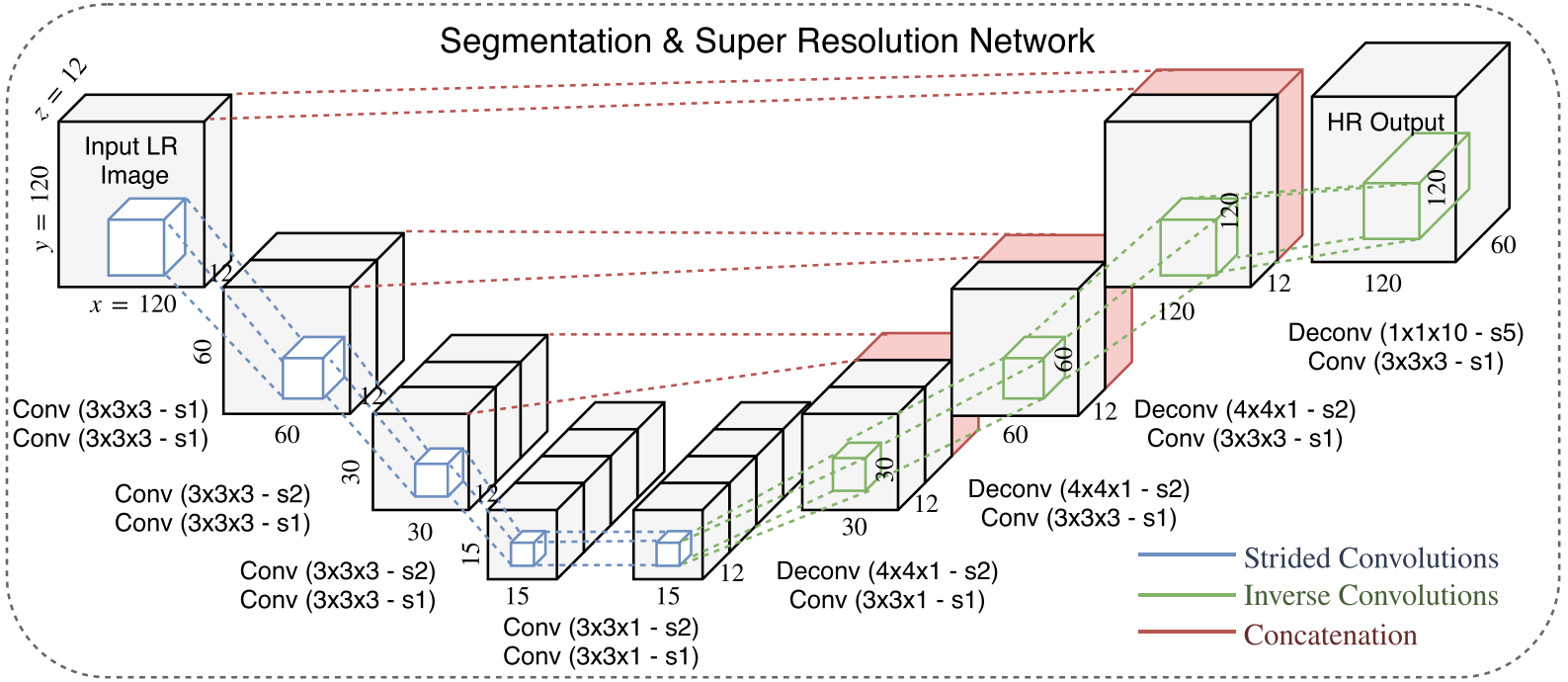
$$L_x = -\sum_{i \in \mathcal{S}} \sum_{c=1}^{C} \log \left(\frac{e^{f_{(c,i)}}}{\sum_j e^{f_{(j,i)}}} \right)$$

- Standard loss for **super-resolution**: L2 or L1 loss

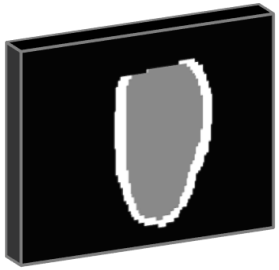
$$\sum_{i \in \mathcal{S}} \| \Phi(\boldsymbol{x}_i, \theta_r) - \boldsymbol{y}_i \|^2$$

Anatomically constrained CNNs

Low-resolution input

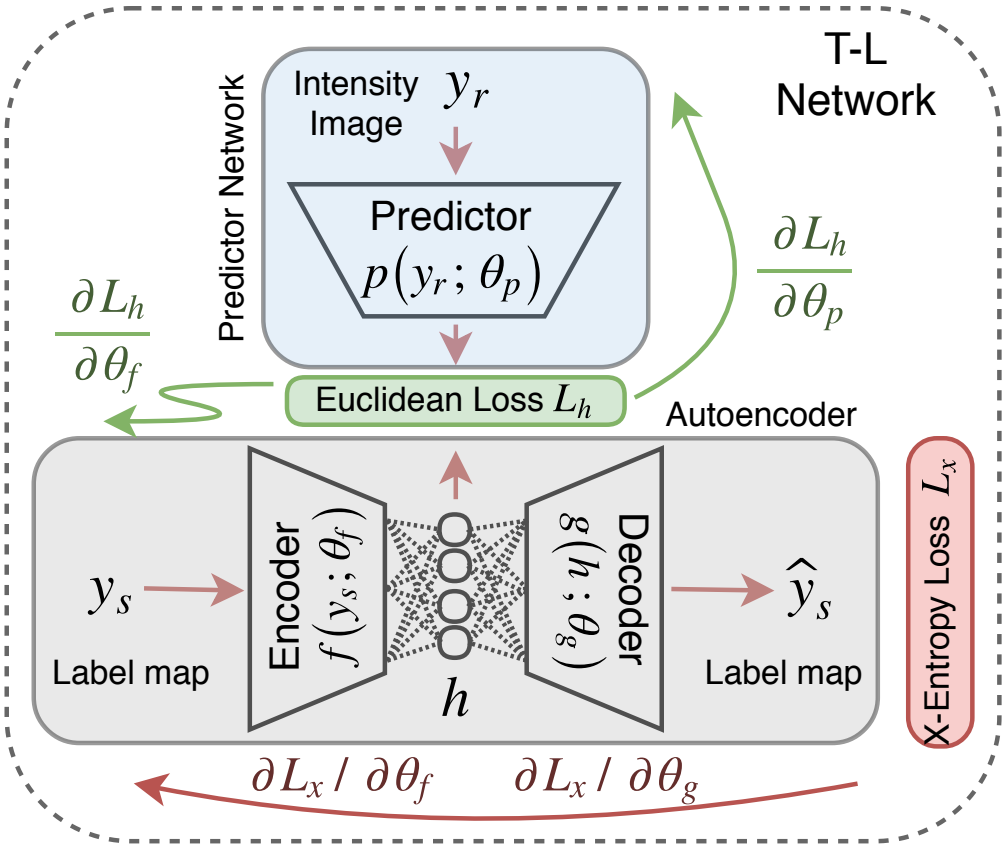


High-resolution output



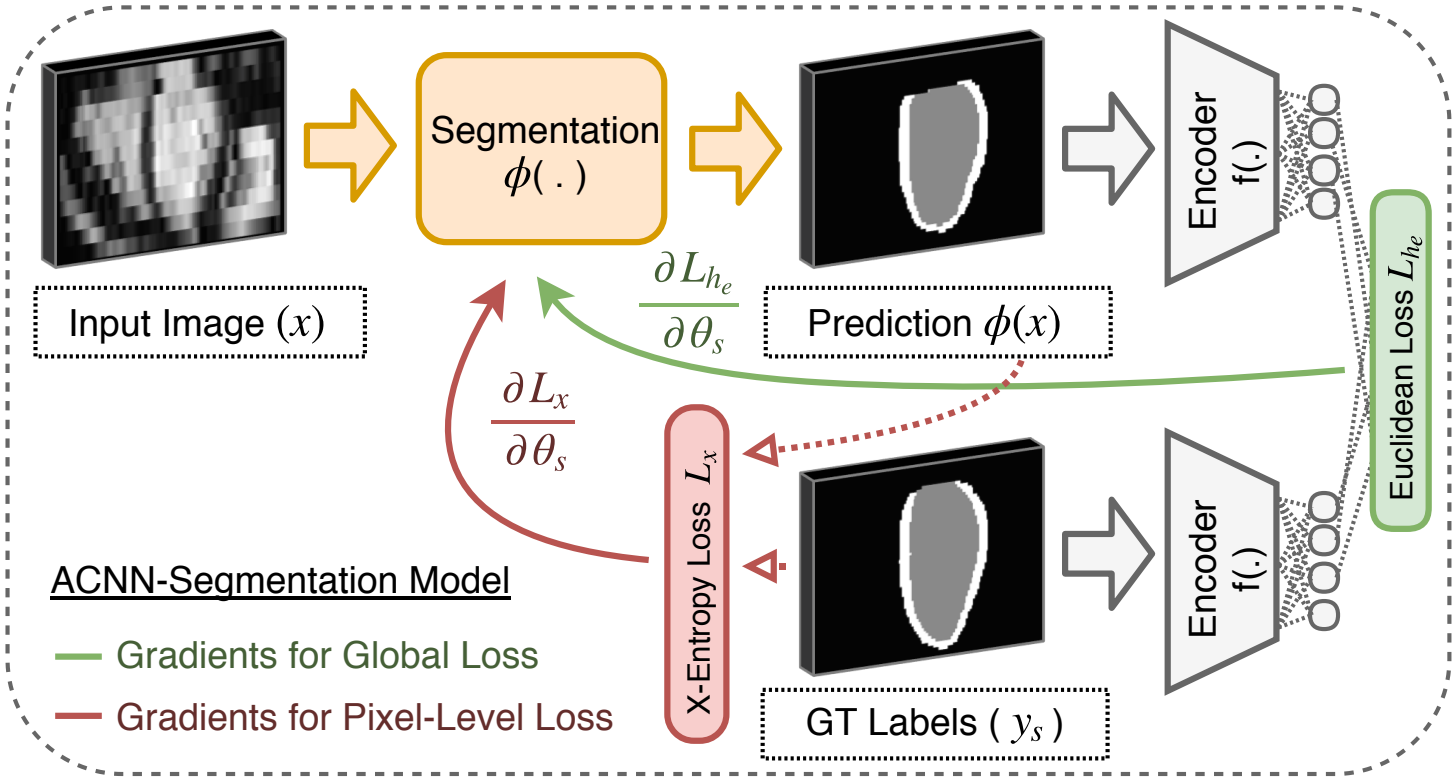
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: T-L networks for representing priors



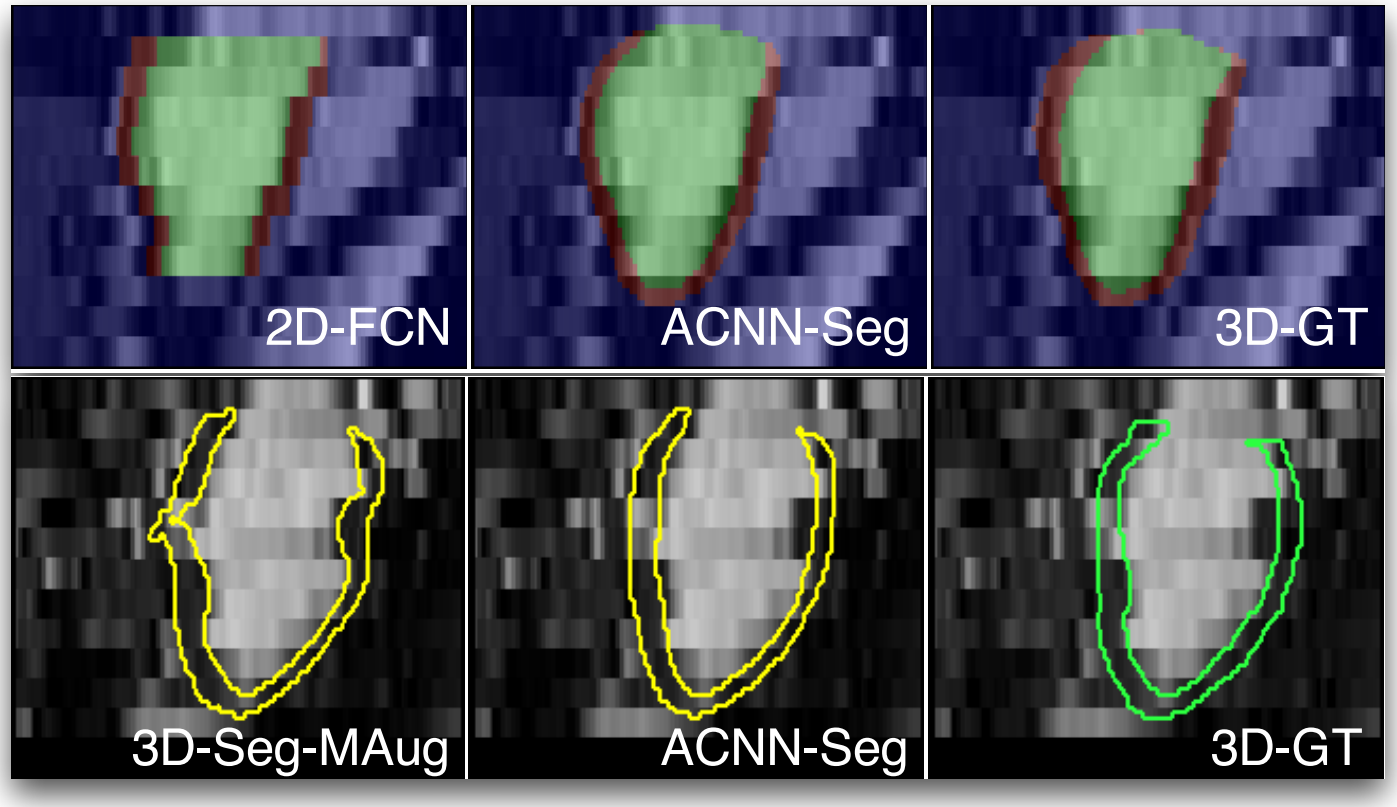
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Segmentation framework



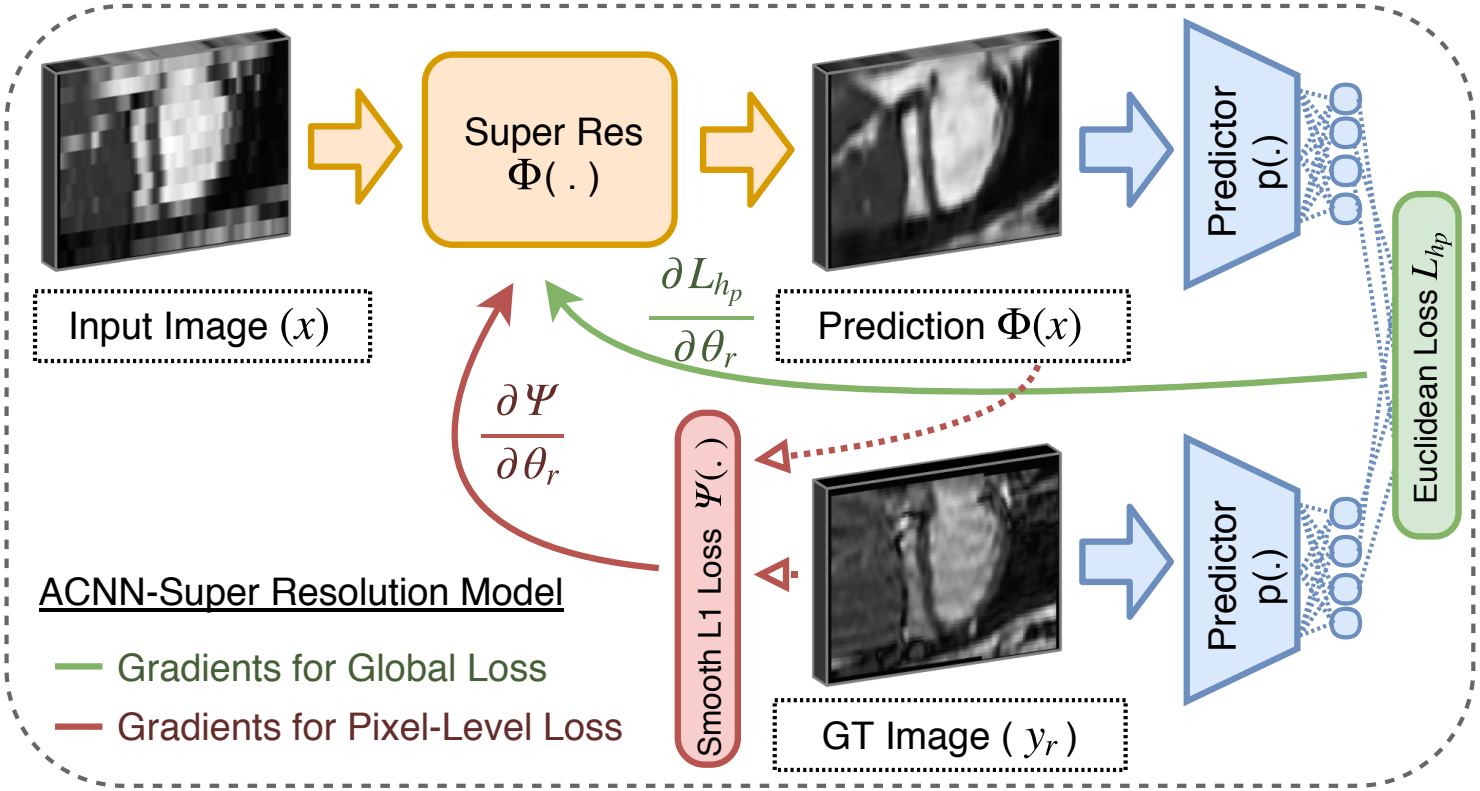
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Segmentation results



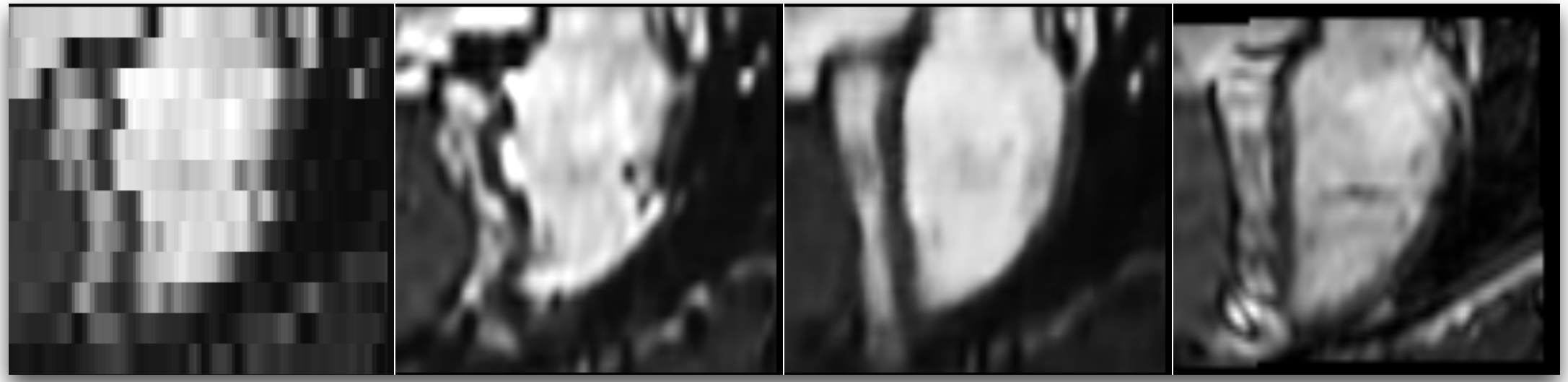
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Super-resolution framework



O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Super-resolution results



Original LR image

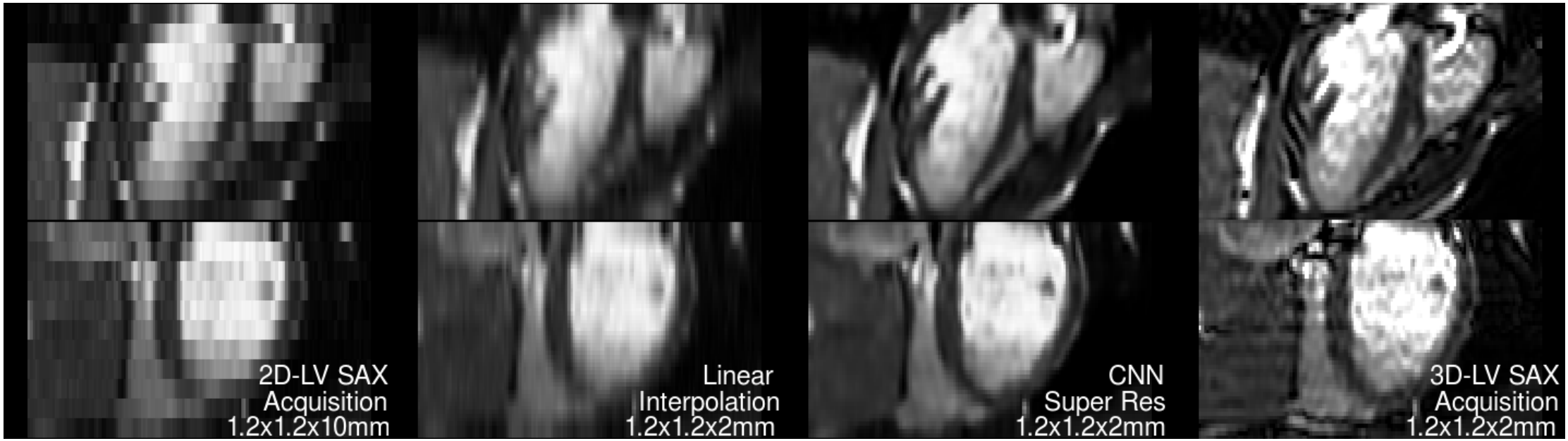
Baseline SR approach

Anatomically constrained SR model

Ground-truth HR image

O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Super-resolution results



O. Oktay et al. IEEE TMI 2017

Overview

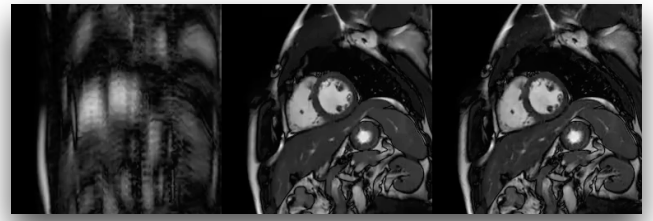


Image reconstruction

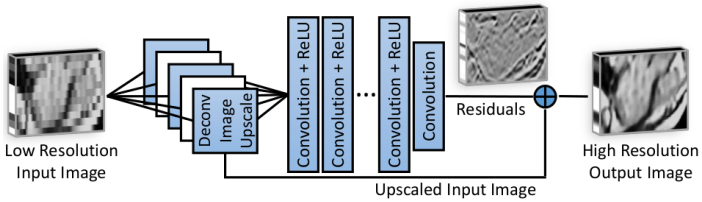


Image super-resolution

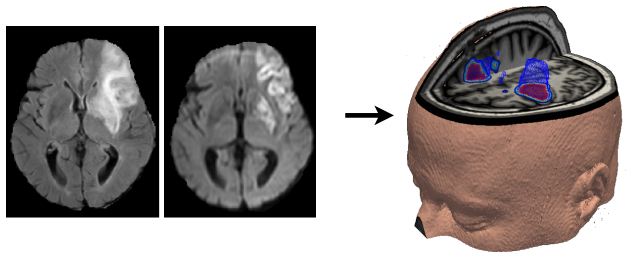


Image segmentation

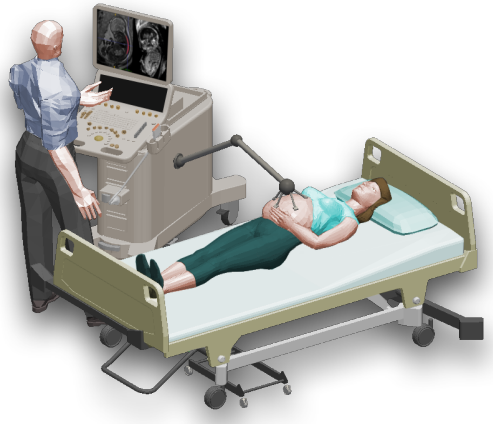
Abdominal View
Confidence: 98%

Image classification

iFIND

intelligent Fetal Imaging and Diagnosis

- Aim: Better fetal screening with US
  - Use multiple US probes to acquire more comprehensive imaging data
  - Use robotic control of US probes to ensure wider field of view
  - Use machine learning for automated US image acquisition and interpretation



- Partners:

wellcome trust PHILIPS EPSRC

KING'S College LONDON Imperial College London UNIVERSITÀ DEGLI STUDI FIRENZE SickKids

Automatic Standard Scan Plane Detection

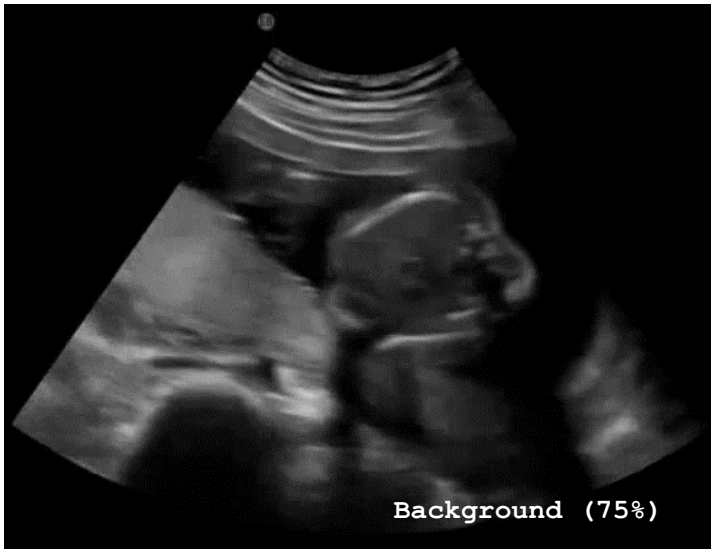
Abdominal View
Confidence: 98%

Lips View
Confidence: 96%

Goal: Do this in real-time on images straight from US machine

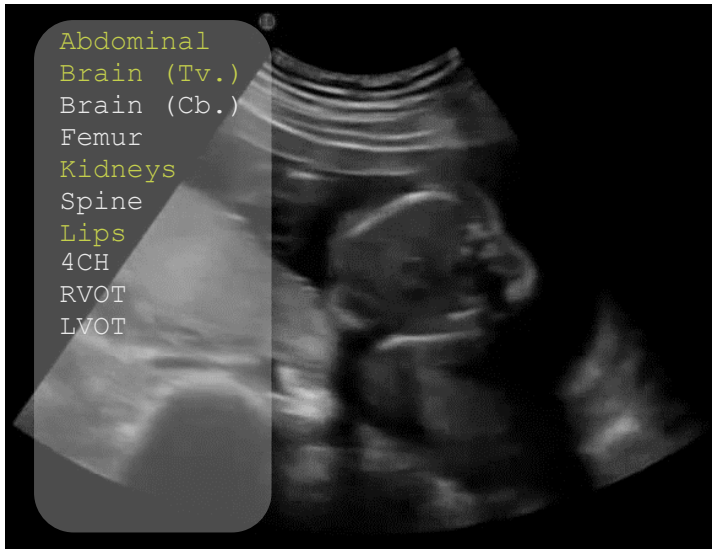
Automatic Standard Scan Plane Detection

- Potential applications:
  - Guidance: Assist inexperienced sonographers



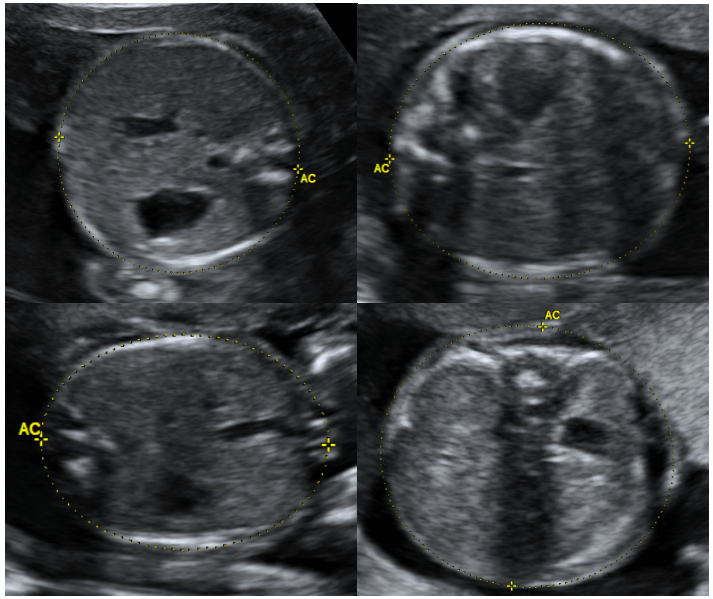
Automatic Standard Scan Plane Detection

- Potential applications:
  - Guidance: Assist inexperienced sonographers
  - Convenience: Automatically make a check list of visited planes



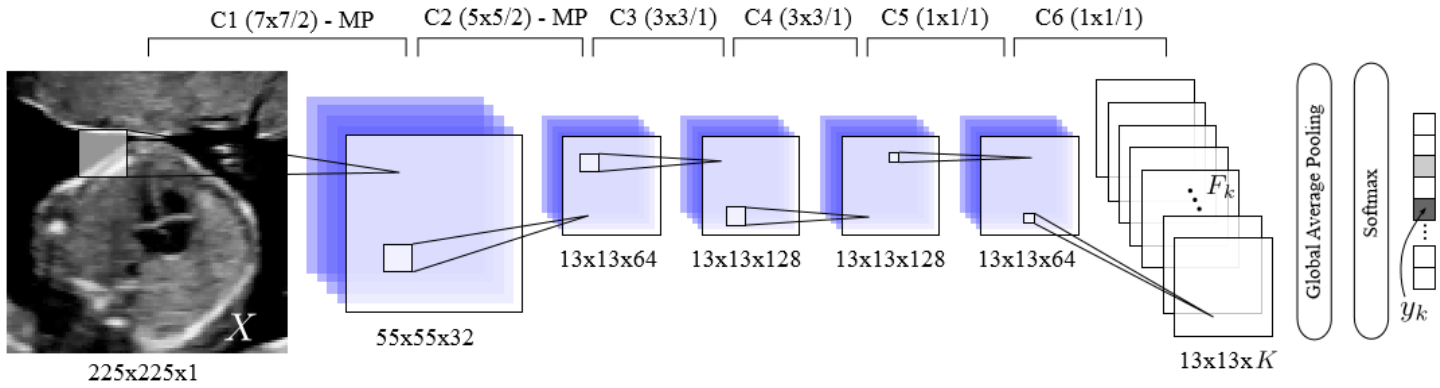
Automatic Standard Scan Plane Detection

- Potential applications:
  - Guidance: Assist inexperienced sonographers
  - Convenience: Automatically make a check list of visited planes
  - Reproducibility: Reduce variability between operators



Automatic Standard Scan Plane Detection: Method

- Fully convolutional neural network:

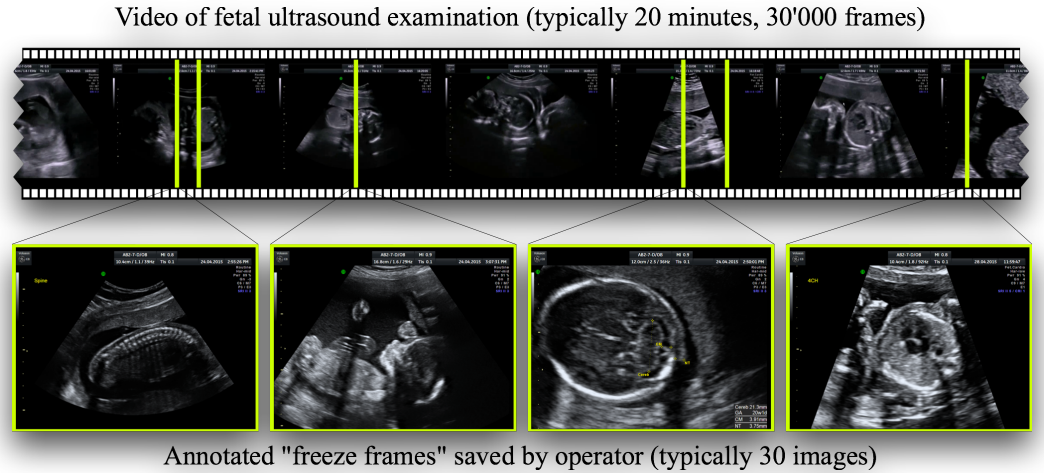


  - Very fast
  - Very accurate

C. Baumgartner et al. MICCAI 2016, IEEE-TMI 2017

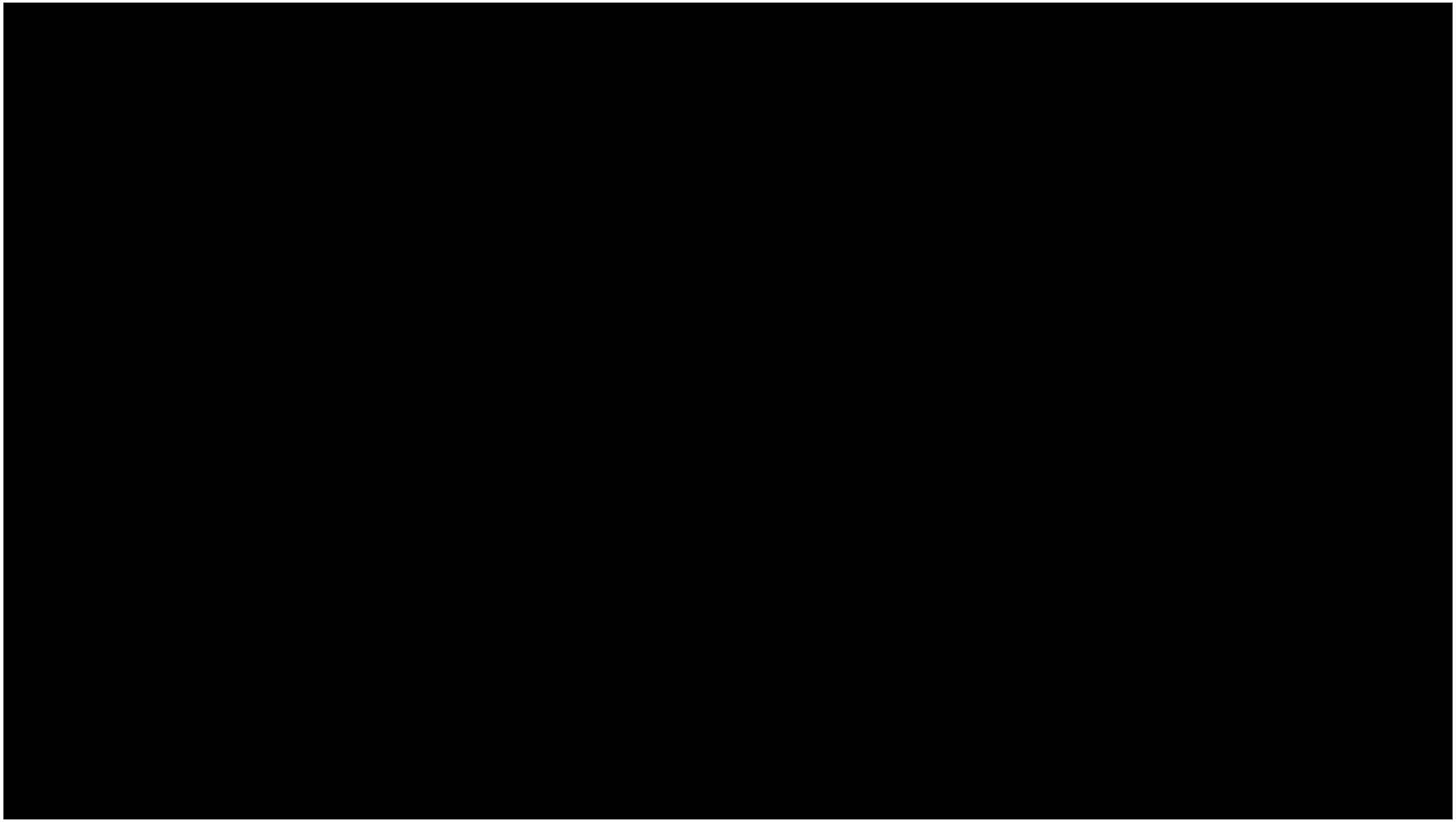
Automatic Standard Scan Plane Detection: Data

- We use very large 2D ultrasound dataset consisting of ***images*** of standard views and ***videos***



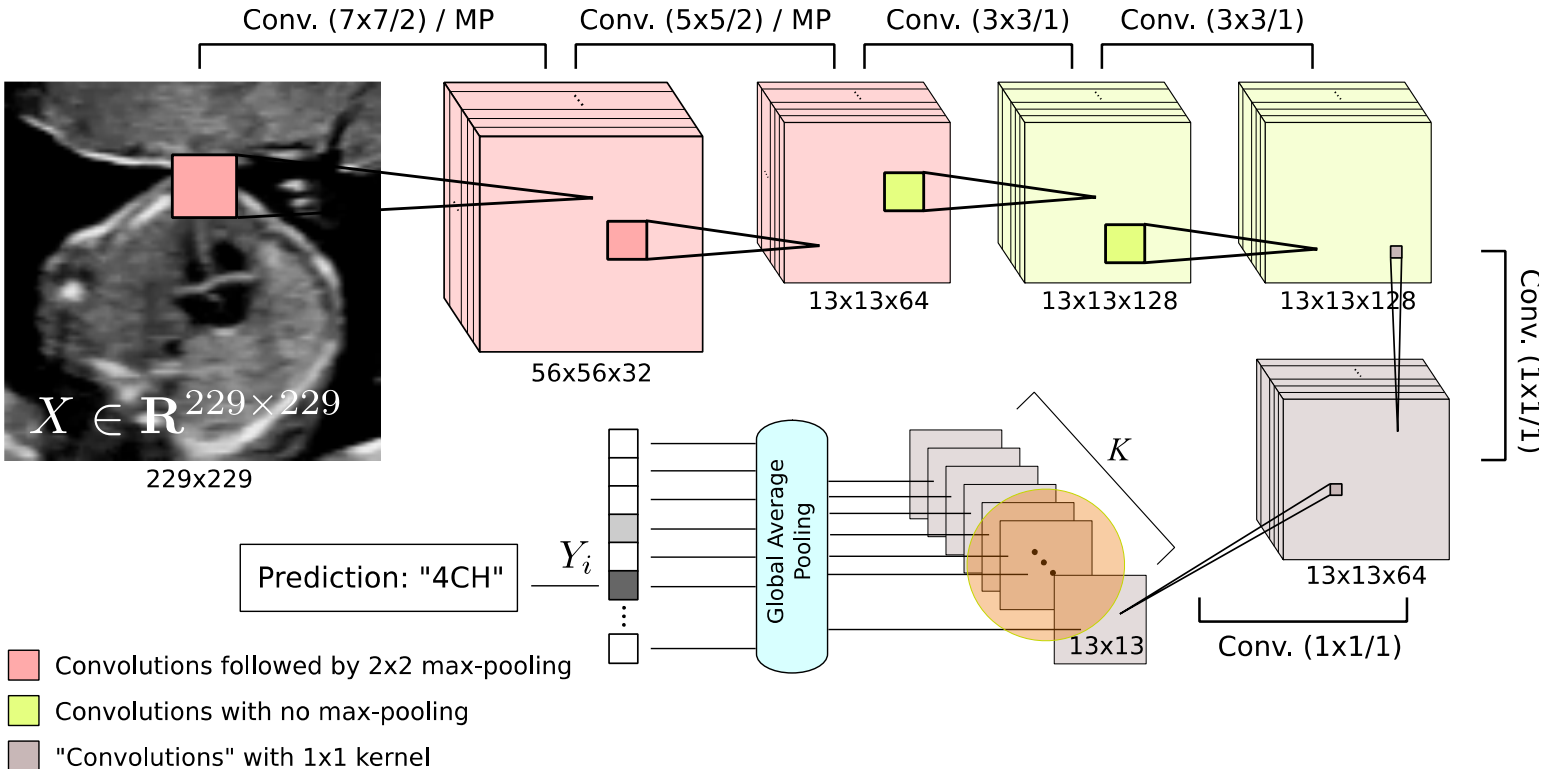
- Data from
  - 2700 patients
  - Between 1200 and 4800 images for each standard plane

Demo



C. Baumgartner et al. MICCAI 2016, IEEE-TMI 2017

Automatic Standard Scan Plane Detection: Localisation

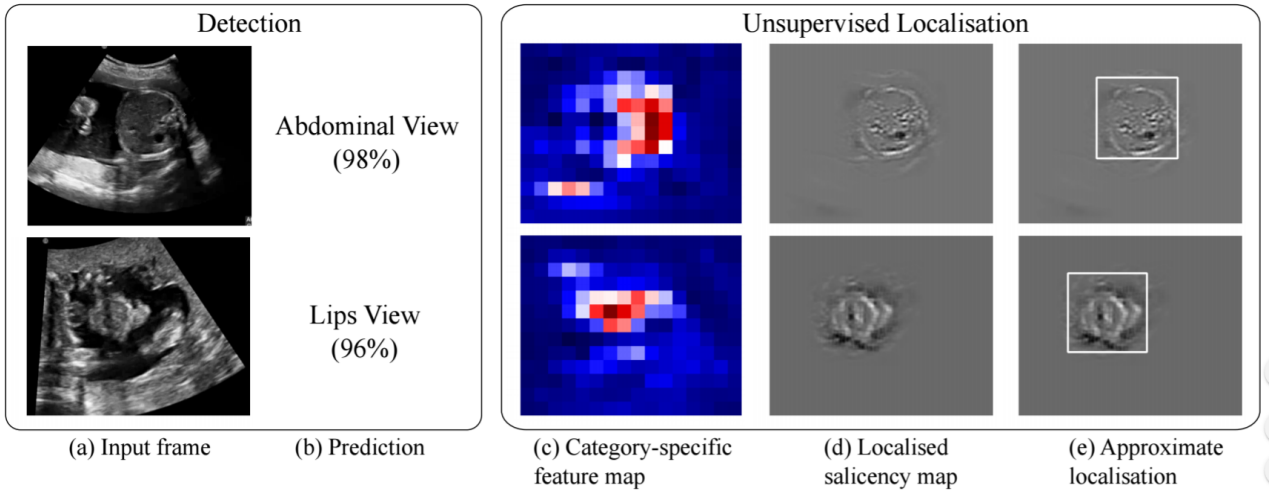


Localisation is (almost) for free in this framework!

C. Baumgartner et al. MICCAI 2016, IEEE-TMI 2017

Automatic Standard Scan Plane Detection: Localisation

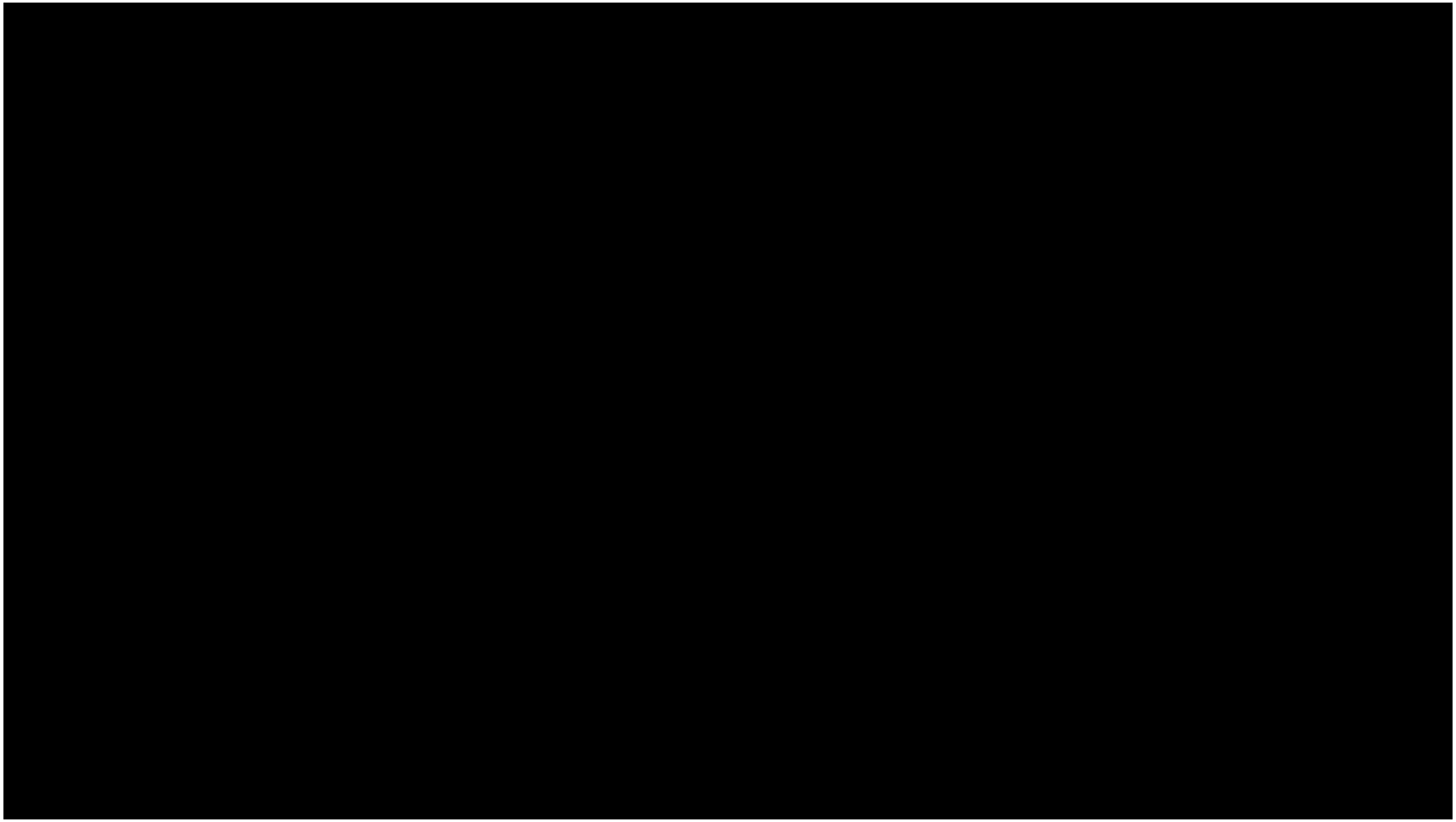
- Can also identify which regions of a frame caused it to make a particular prediction



- This can be used for localisation of the fetal anatomy without having bounding boxes for training

C. Baumgartner et al. MICCAI 2016, IEEE-TMI 2017

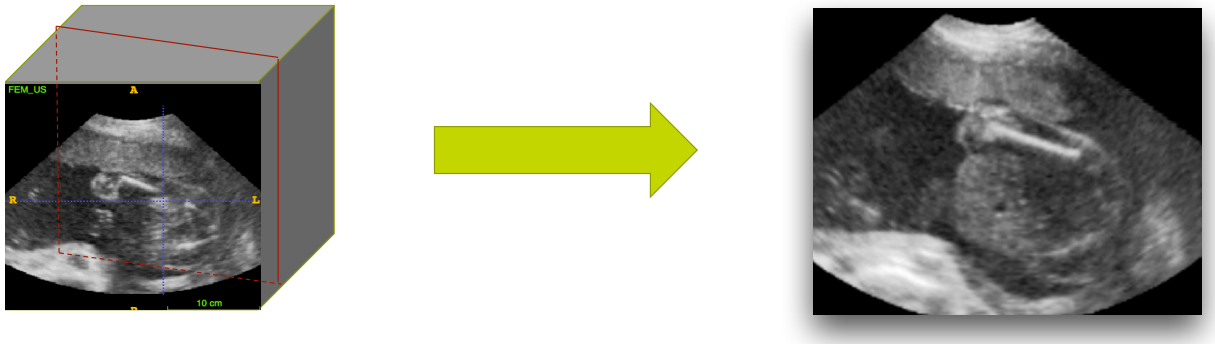
Demo



C. Baumgartner et al. MICCAI 2016, IEEE-TMI 2017

Standard Plane Detection from 3D Ultrasound

- Motivation
  - 3D ultrasound data is hard to interpret directly on the US scanner
  - We aim for a system that can automatically extract standard 2D views from a 3D view at any probe position
- Eventually, we would like this to work in real-time



Overview

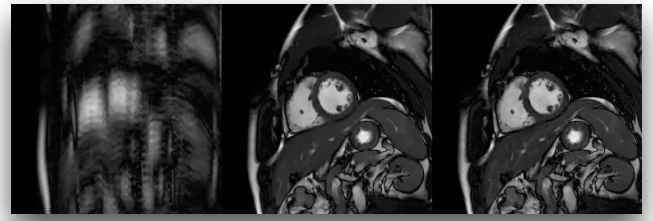


Image reconstruction

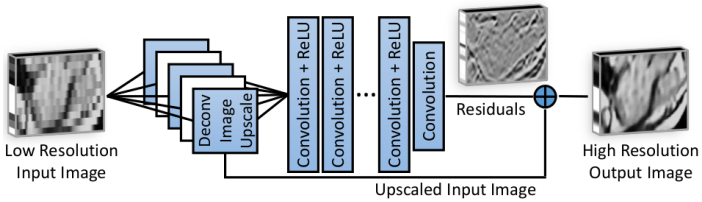


Image super-resolution

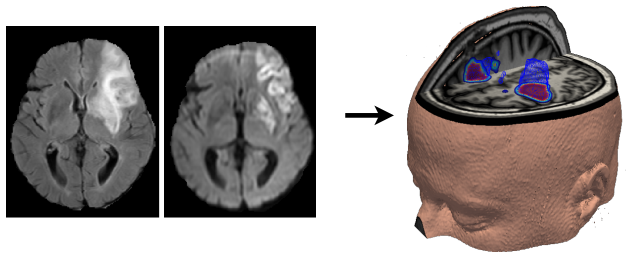
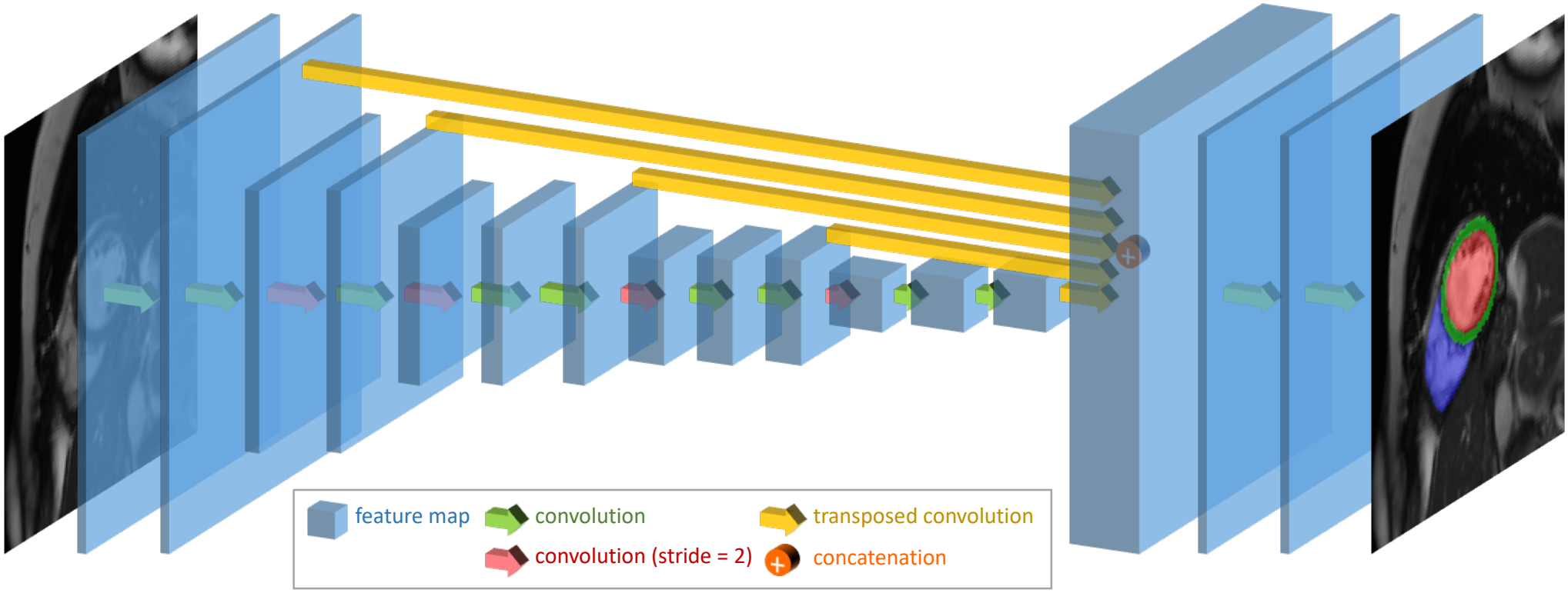


Image segmentation

Abdominal View
Confidence: 98%

Image classification

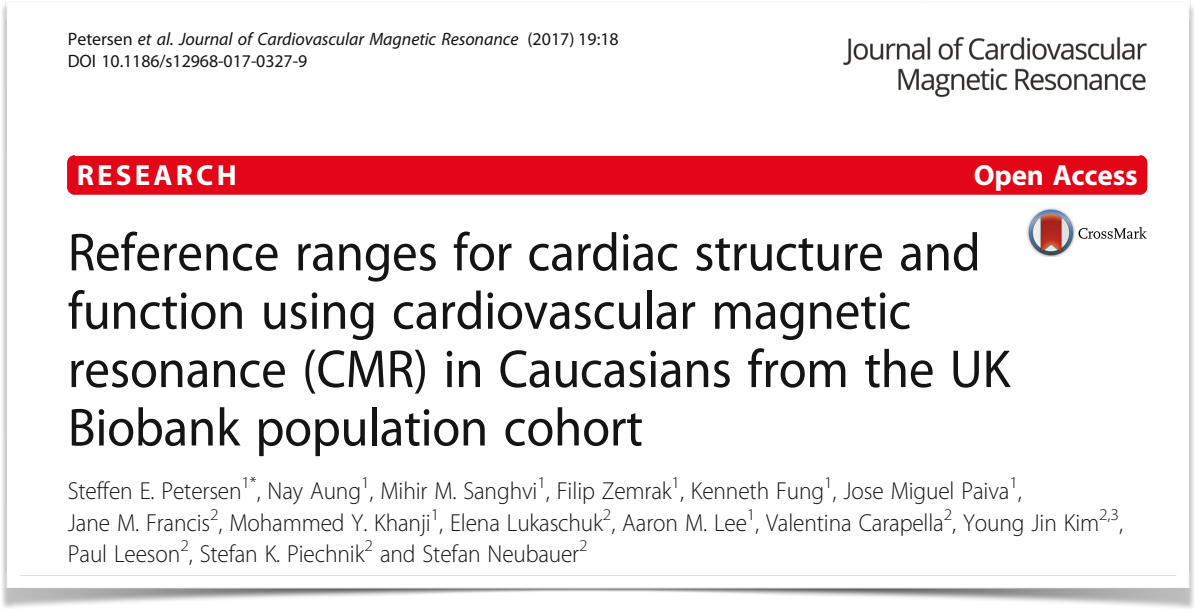
Convolutional Neural Networks for Medical Image Segmentation



W. Bai et. JCMR, in press 2018
arXiv:1710.09289v3

Image segmentation as a machine learning problem

- Manual annotations of **4,872 subjects** (QMUL/Oxford) with **93,128 pixelwise annotated 2D images** slices
- Divided into training/validation/test: 3,972/300/600



Petersen *et al. Journal of Cardiovascular Magnetic Resonance* (2017) 19:18
DOI 10.1186/s12968-017-0327-9

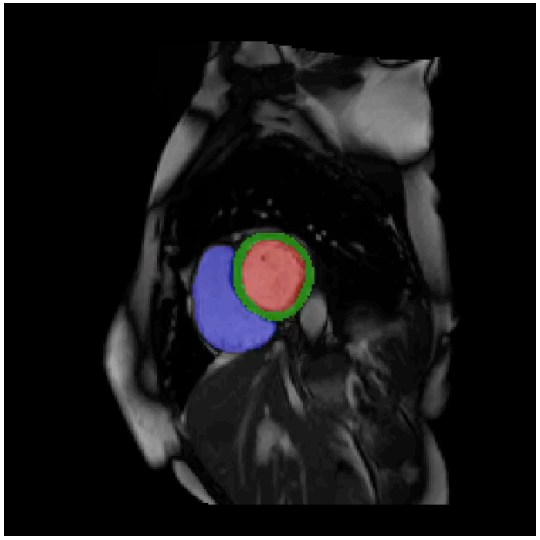
Journal of Cardiovascular Magnetic Resonance

RESEARCH Open Access

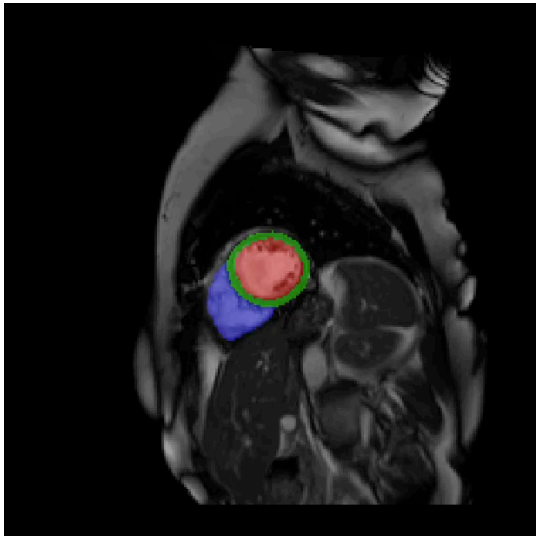
Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort

Steffen E. Petersen[1*], Nay Aung[1], Mihir M. Sanghvi[1], Filip Zemrak[1], Kenneth Fung[1], Jose Miguel Paiva[1], Jane M. Francis[2], Mohammed Y. Khanji[1], Elena Lukaschuk[2], Aaron M. Lee[1], Valentina Carapella[2], Young Jin Kim[2,3], Paul Leeson[2], Stefan K. Piechnik[2] and Stefan Neubauer[2]

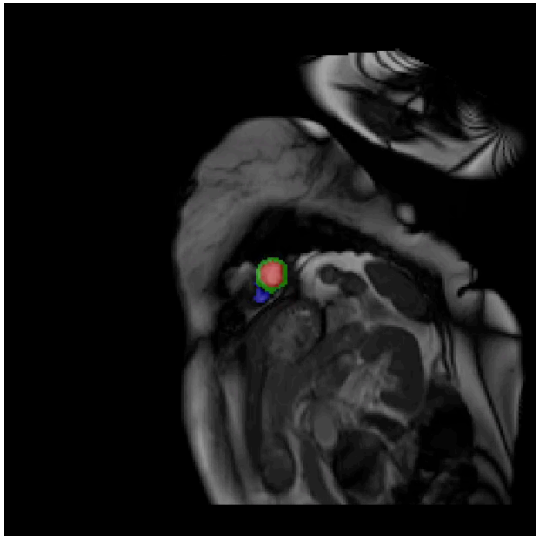
W. Bai et. JCMR, in press 2018
arXiv:1710.09289v3



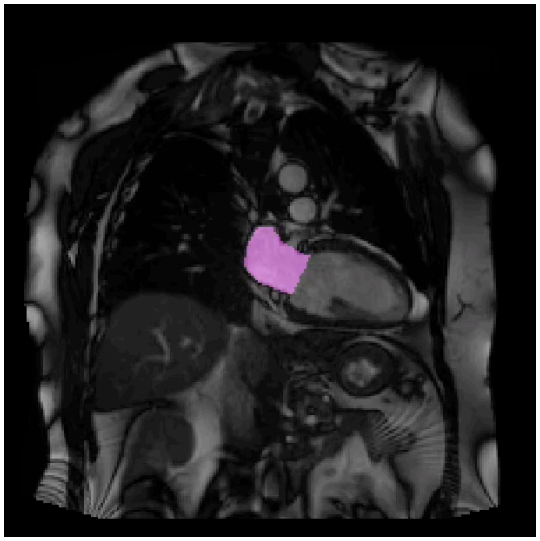
SA, basal



SA, mid-ventricular



SA, apical



LA, 2 chamber

LA, 4 chamber

W. Bai et. JCMR, in press 2018
arXiv:1710.09289v3

Evaluation of segmentation accuracy
Comparison to expert observers

(a) Absolute difference

	Auto vs Man (n = 600)	O1 vs O2 (n = 50)	O2 vs O3 (n = 50)	O3 vs
LVEDV (mL)	$6.1_{\pm 5.3}$	$6.1_{\pm 4.4}$	$8.8_{\pm$	
LVESV (mL)	$5.3_{\pm 4.9}$	$4.1_{\pm 4.2}$		
LVM (gram)	$6.9_{\pm 5.5}$	$4.2_{\pm$		± 4.8
RVEDV (mL)	$8.5_{\pm 7.1}$			$8.7_{\pm 5.8}$
RVESV (mL)	$7.2_{\pm 6.8}$		± 5.5	$11.7_{\pm 6.9}$

ice

		1 vs O2 (n = 50)	O2 vs O3 (n = 50)	O3 vs O1 (n = 50)
	± 3.5	$4.2_{\pm 3.1}$	$6.3_{\pm 3.3}$	$3.4_{\pm 2.2}$
	$9.5_{\pm 9.5}$	$6.8_{\pm 7.5}$	$12.5_{\pm 8.5}$	$11.7_{\pm 5.1}$
	$8.3_{\pm 7.6}$	$4.4_{\pm 3.3}$	$6.0_{\pm 3.7}$	$6.7_{\pm 4.6}$
EDV (%)	$5.6_{\pm 4.6}$	$8.0_{\pm 5.0}$	$4.2_{\pm 3.1}$	$5.7_{\pm 3.6}$
RVESV (%)	$11.8_{\pm 12.2}$	$30.6_{\pm 15.5}$	$10.9_{\pm 8.3}$	$16.9_{\pm 9.2}$

Computer performs as well as different expert observers

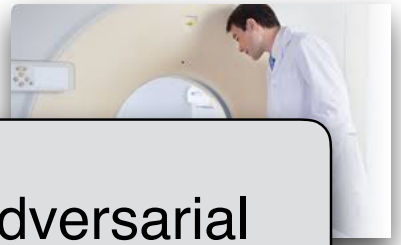
Automated

Manual

W. Bai et. submitted to JCMR, 2018
arXiv:1710.09289v3

Challenges for medical image segmentation: Deployment in the clinic

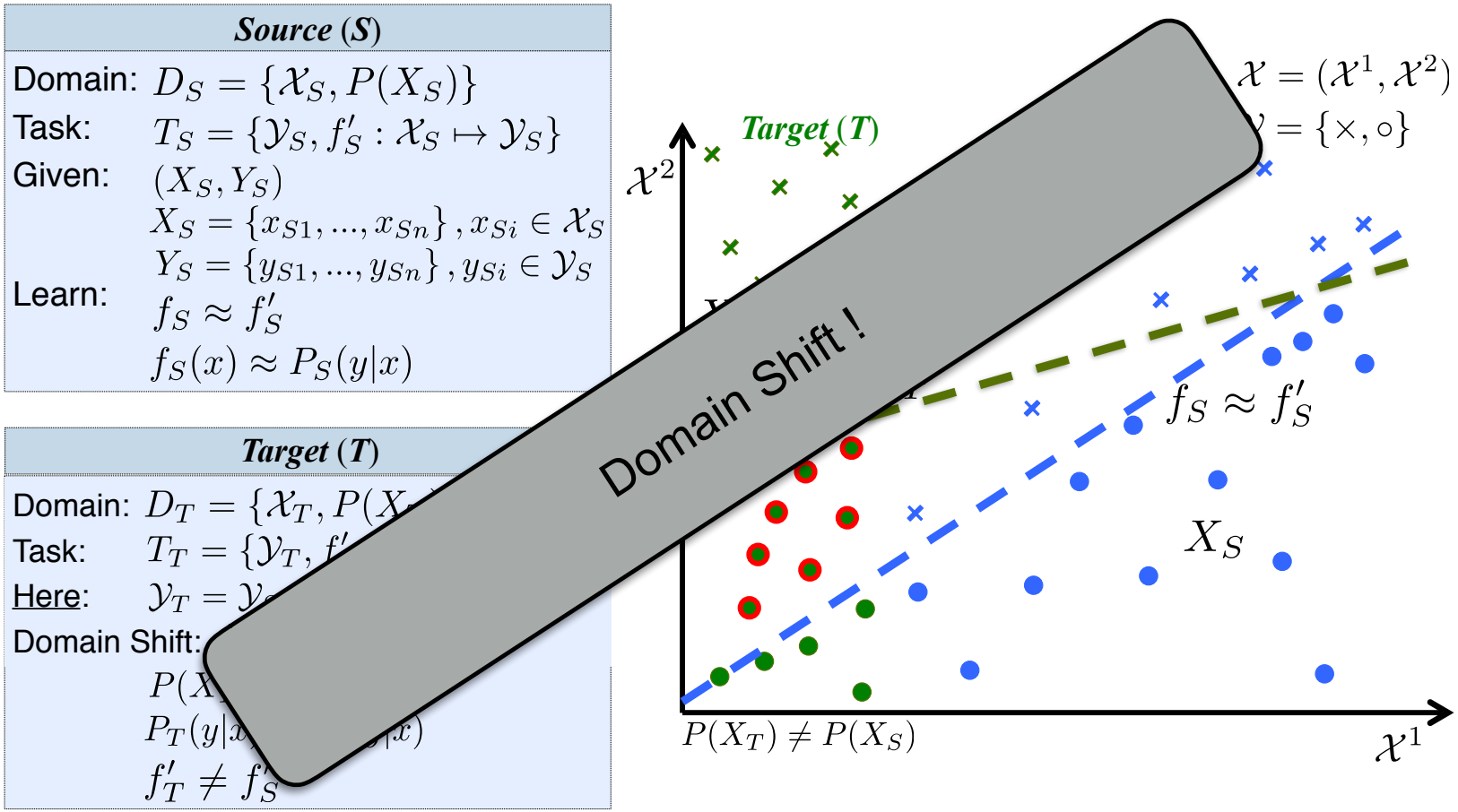
- ML-based segmentation often degrades when deployed in clinical scenarios
- This is
  training
  variati
  - scar
  - scar
- Manua
  test do



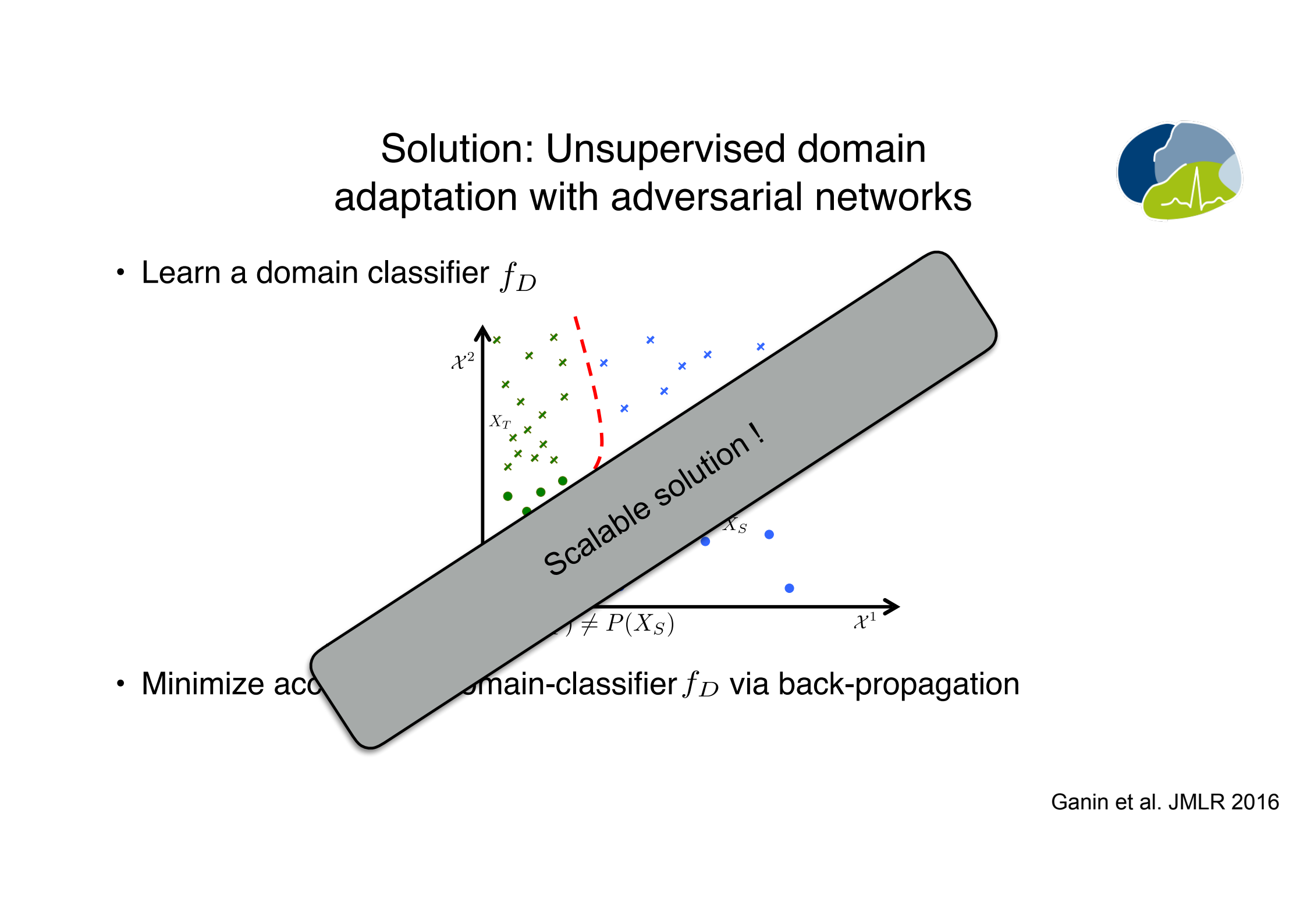
Unsupervised domain adaptation using adversarial neural networks can be used to train a CNN-based segmentation

- which is more invariant to differences in the input data
- which does not require any annotations on the test domain

Deploying machine learning into clinical practice: What is the problem?



Solution: Unsupervised domain adaptation with adversarial networks



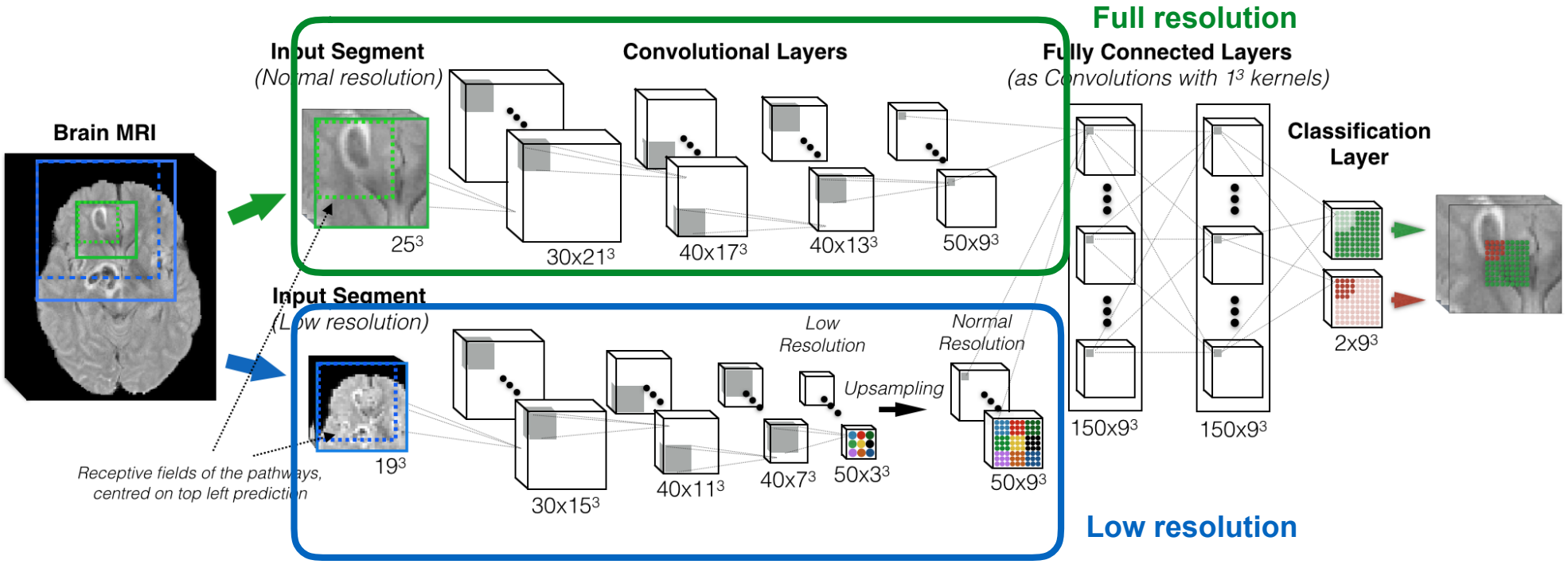
- Learn a domain classifier f_D

- Minimize acc omain-classifier f_D via back-propagation

Ganin et al. JMLR 2016

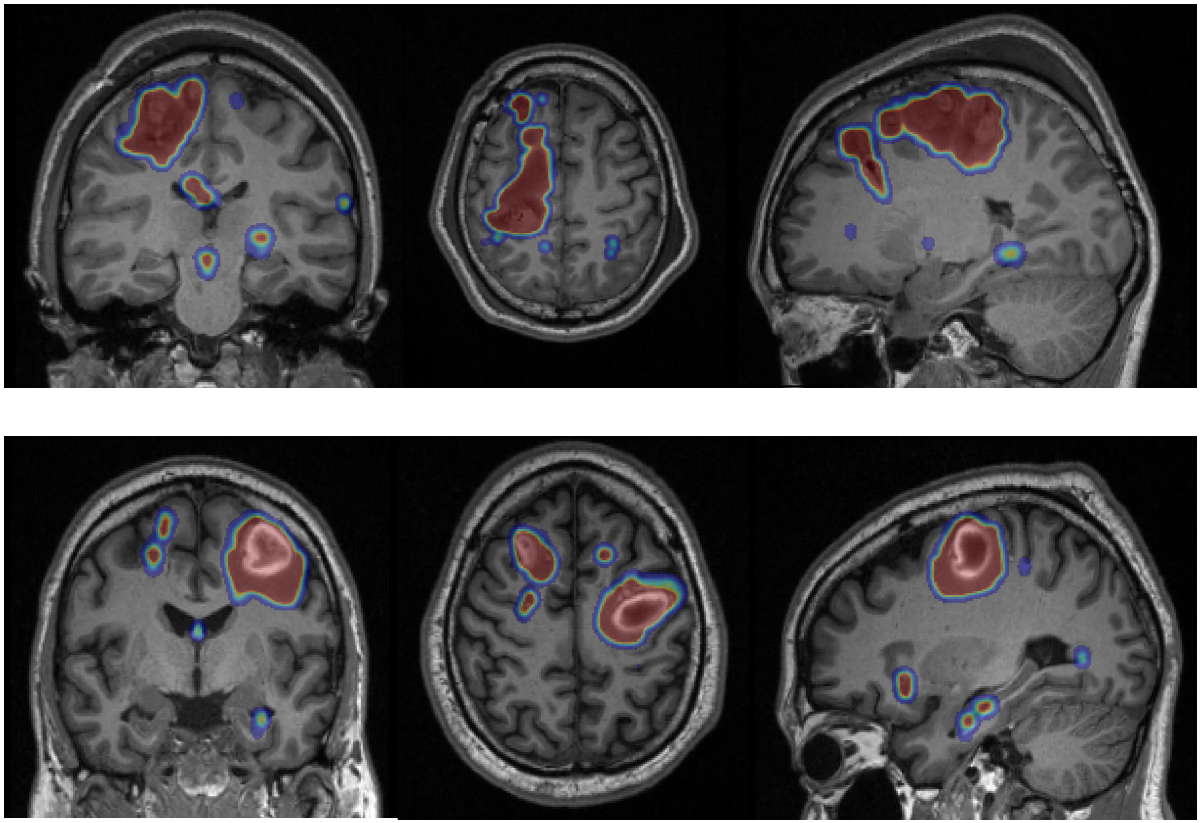
DeepMedic: Overview

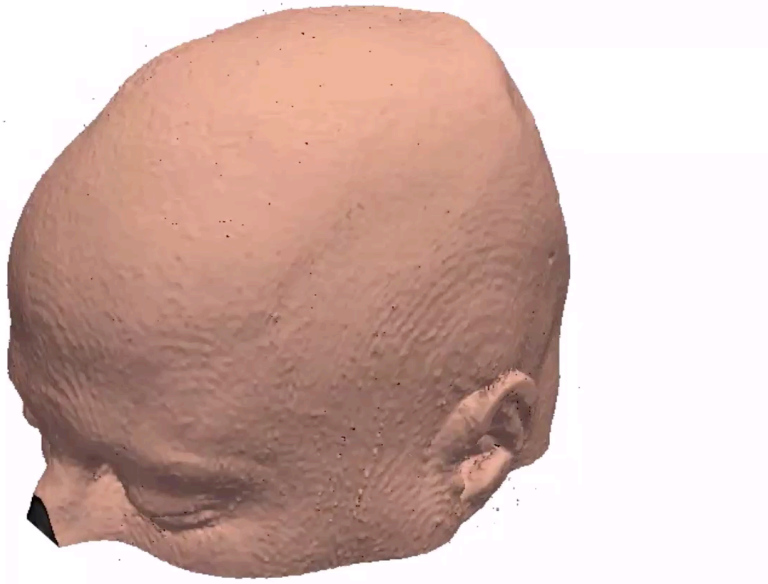
- The size of receptive field of CNNs is important:
  - Large receptive increases computation and memory requirements
  - Pooling leads to loss of the spatial information



K. Kamnitsas et al. Medical Image Analysis, 2016

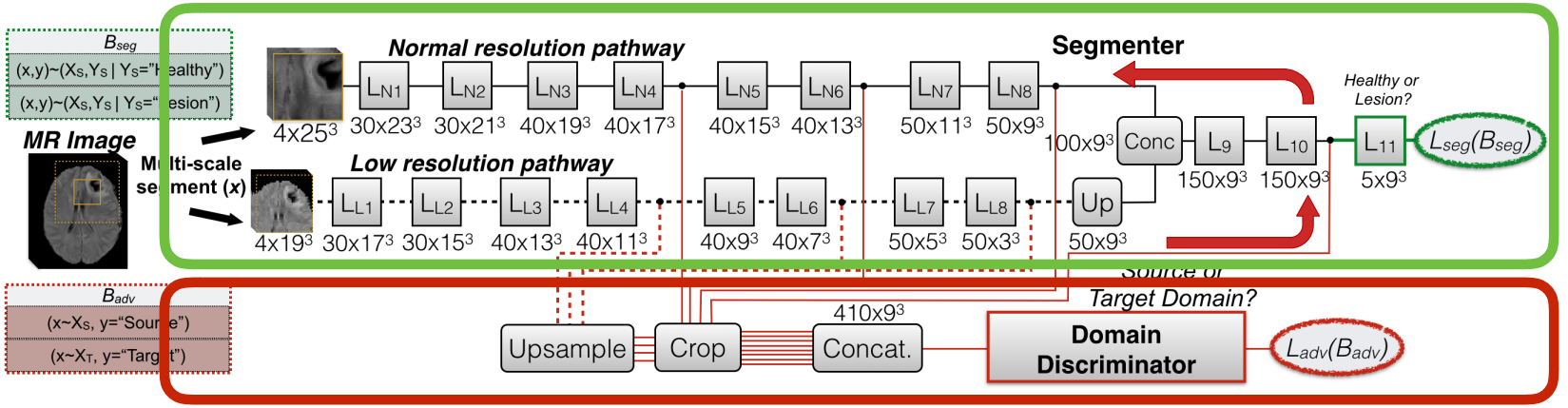
DeepMedic in Action





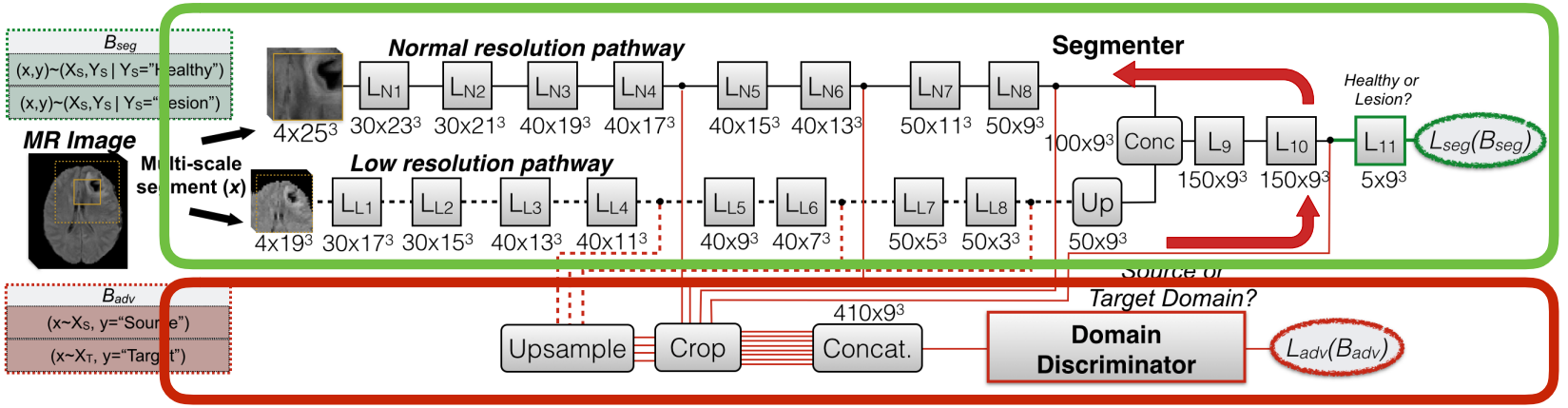
K. Kamnitsas et al. Medical Image Analysis, 2016

DeepMedic: Unsupervised domain adaptation with adversarial networks



K. Kamnitsas et al. IPMI 2017, arXiv:1612.08894

DeepMedic: Unsupervised domain adaptation with adversarial networks

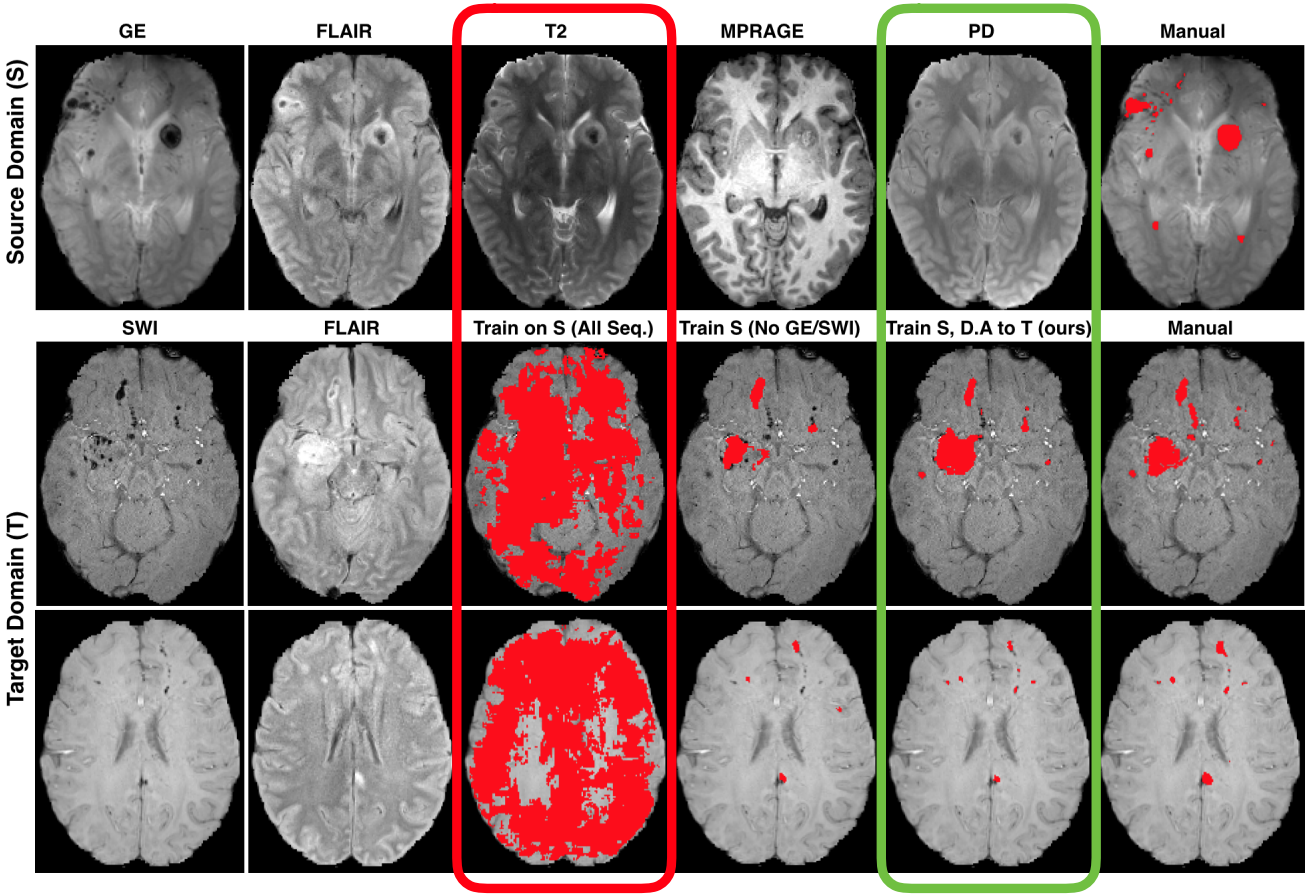


Segmenter: $\mathcal{L}_{seg} = -\frac{1}{m}\sum_{i=1}^{m}[f(x_i) = y_i]log(f(x_i)) \quad , (x_i, y_i) \sim (X_S, Y_S)$

Domain Discr.: $\mathcal{L}_{adv} = -\frac{1}{m}\sum_{i=1}^{m} log(f_D(h(x_i))) - \frac{1}{m}\sum_{j=1}^{m} log(1 - f_D(h(x_j))) \; , x_i \sim X_S \; x_j \sim X_T$

K. Kamnitsas et al. IPMI 2017, arXiv:1612.08894

DeepMedic: Unsupervised domain adaptation with adversarial networks



K. Kamnitsas et al. IPMI 2017, arXiv:1612.08894

Challenges for medical image segmentation: DeepMedic, FCN & U-Net

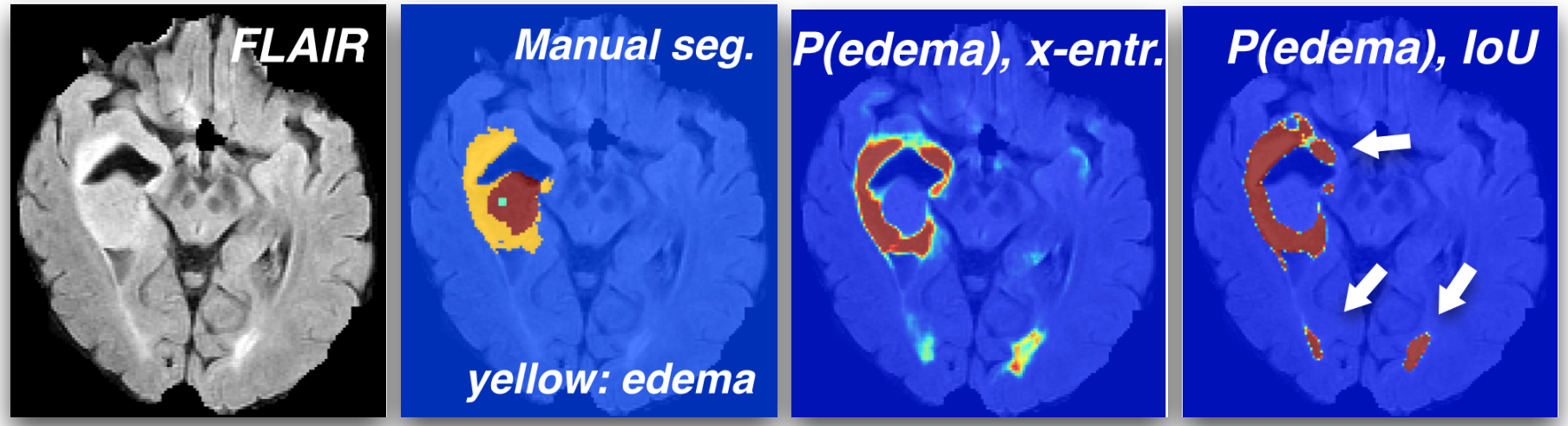
- **The good:**
  - There are some good/promising CNN-based segmentation approaches (DeepM
- **The ba**
  - A lot o
  - Archit
  - Archit
- **The ugly:**
  - Chosen model & config may be suboptimal for other data/task
  - Results and conclusions of analysis are strongly biased

Ensemble of Multiple Models & Architectures (EMMA)

Performance **insensitive** to suboptimal configuration

Behaviour **unbiased** by architecture & configuration

Challenges for medical image segmentation: Behaviour and performance is variable



Model trained with cross-entropy loss

Model trained with IoU (Dice) loss

Ensemble of Multiple Models and Architectures (EMMA)

Need to learn: $P(Y|X)$

Approximate it by model: $P(Y|X;\theta_m,m)$

with learnt parameters $\theta_m = \min_{\theta_m} d\left(P\left(Y|X;\theta_m,m\right), P(Y|X)\right)$, ***d*** the loss.

Model is defined by chosen meta-parameters ***m.***

Commonly *m* is neglected, but it biases the results!

We define stochastic random variable ***M***, over configurations of interest.

Need to marginalise out influence of ***M***:

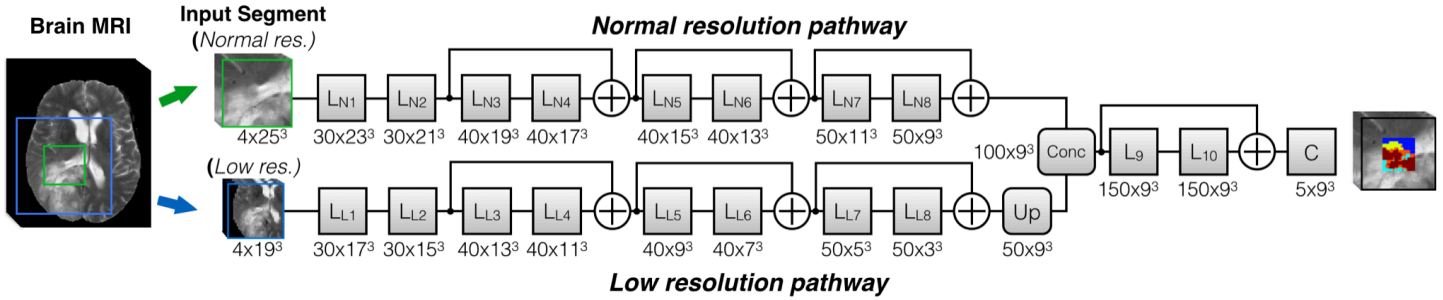
$$P(Y|X) = \sum_{\forall m \in M} P(Y, M=m|X) = \sum_{\forall m \in M} P(Y|X, M=m)\, P(M=m)$$

EMMA approximate the joint by ensembling individual models:

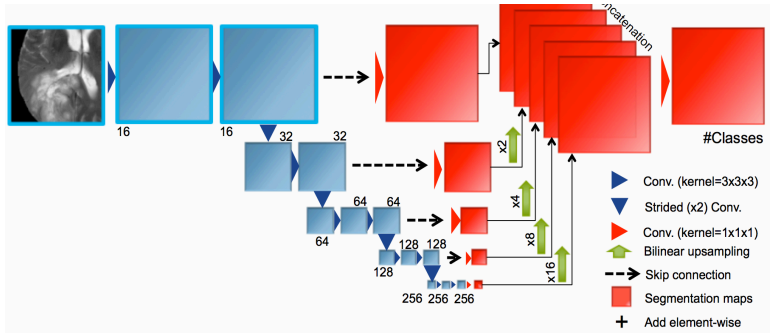
$$\sim P_{EMMA}(Y|X) = \sum_{\forall m \in M} P(Y|X;\theta_m,m)\, \frac{1}{|M|}$$

M: Network architectures

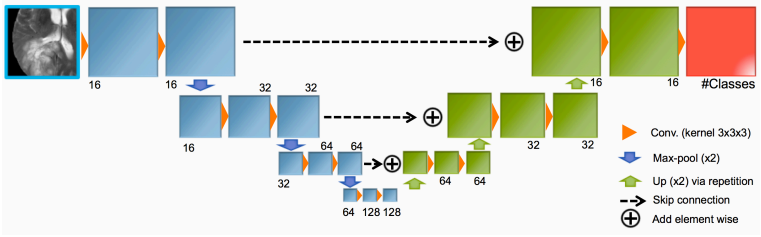
DeepMedic [Kamnitsas 2015, 2016, 2017]



FCN [Long 2015]:



U-Net [Ronneberger 2015]:



M: Network configurations

- Architecture configuration:
  - depth, width, scales, residuals, etc.
- Training Loss:
  - Cross-Entropy, IoU, DSC, etc.
- Sampling strategy:
  - equally per class, foreground/background, etc.
- Optimisation:
  - optimizer, learning rate, momentum, regulariser…
- Data normalisation:
  - z-score, bias field correction, histogram matching


BraTS
Multimodal Brain Tumor Segmentation Challenge
Led by CBICA
1st Place
2017 MICCAI BraTS Challenge
(Segmentation Task)
K. Kamnitsas, et al. "Ensembles of Multiple Models and
Architectures for Robust Brain Tumour Segmentation"

BRATS'17 Challenge: Quantitative validation

- **EMMA: 2 x DeepMedic, 3 x FCNs, 1 x U-Net**
  - Different training losses, sampling strategies, widths, depths, configurations
  - No config was heavily optimised for the task (3/6 nets were quite suboptimal)

	DSC			Sensitivity			Specificity			Hausdorff_95		
	Enh.	Whole	Core	Enh.	Whole	Core	Enh	Whole	Core	Enh.	Whole	Core
EMMA	**75.7**	**90.2**	82.0	**79.0**	90.9	78.3	99.8	99.5	**99.9**	**4.22**	4.56	**6.11**
UCL-TIG	75.2	89.7	**82.5**	77.1	**91.2**	**83.9**	99.8	99.4	99.7	4.78	**3.97**	7.60
MIC_DKFZ	73.2	89.6	79.7	**79.0**	89.6	78.1	99.8	**99.6**	**99.9**	4.55	6.97	9.48

- **Robustness:**
  - EMMA of all 6 was better than individuals.
  - Ensemble of 3 best nets was only marginally better than EMMA of all 6 nets.

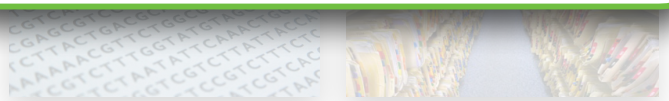
Summary and Conclusions

- Deep learning already plays a crucial role in medical imaging for
  - Image a
  - Image c
- Applications of deep learning in computer-aided decision support have been limited so far
  - But there is some (unjus
- There is great pote
  truly intelligent computer-aided diagnosis
  - Learning from unlabelled, large-scale population data
  - Integration of imaging and non-imaging informat
    clinical records and genetics

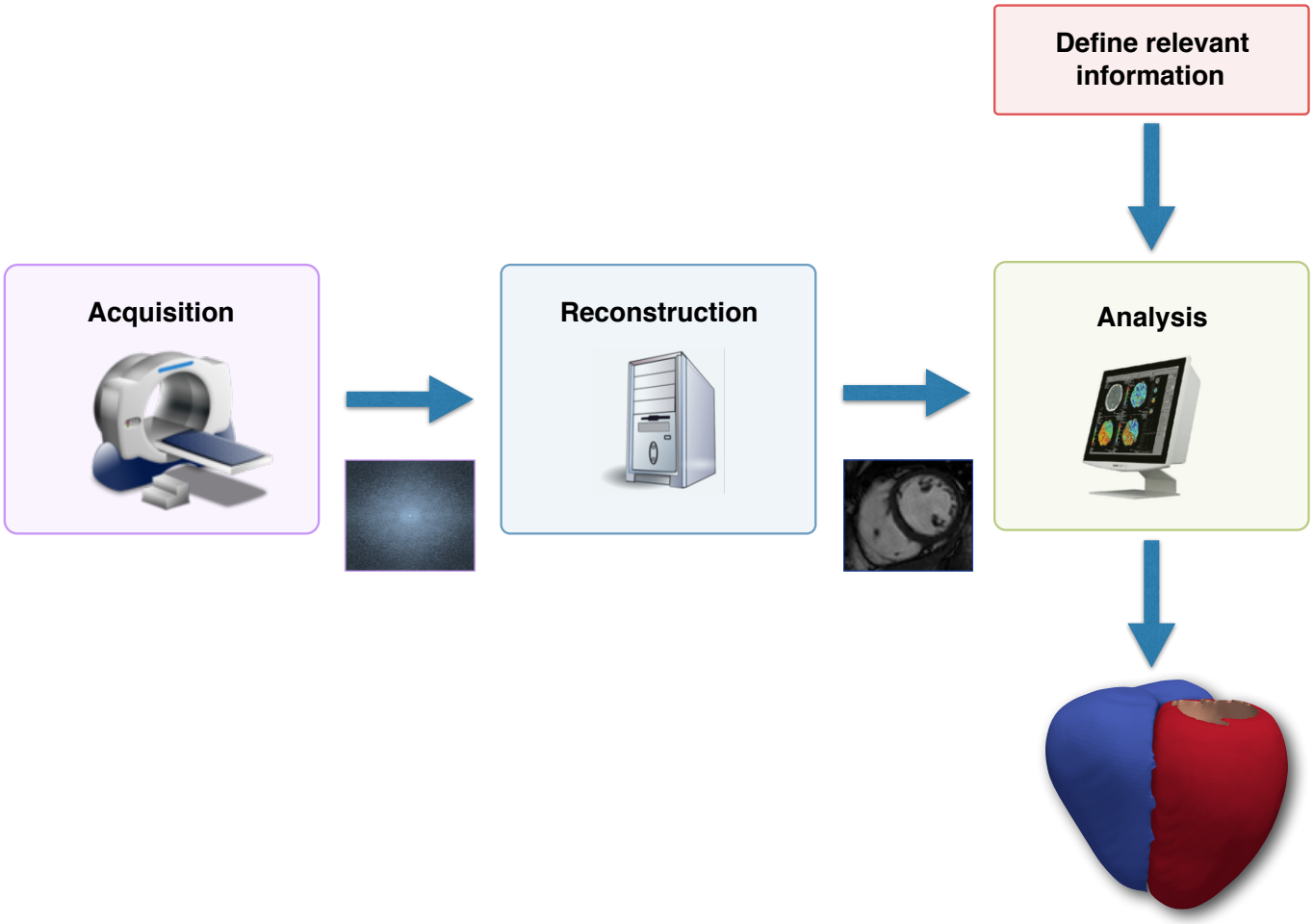
Validation is challenging

Requires collaboration between computer scientists, engineers and clinicians

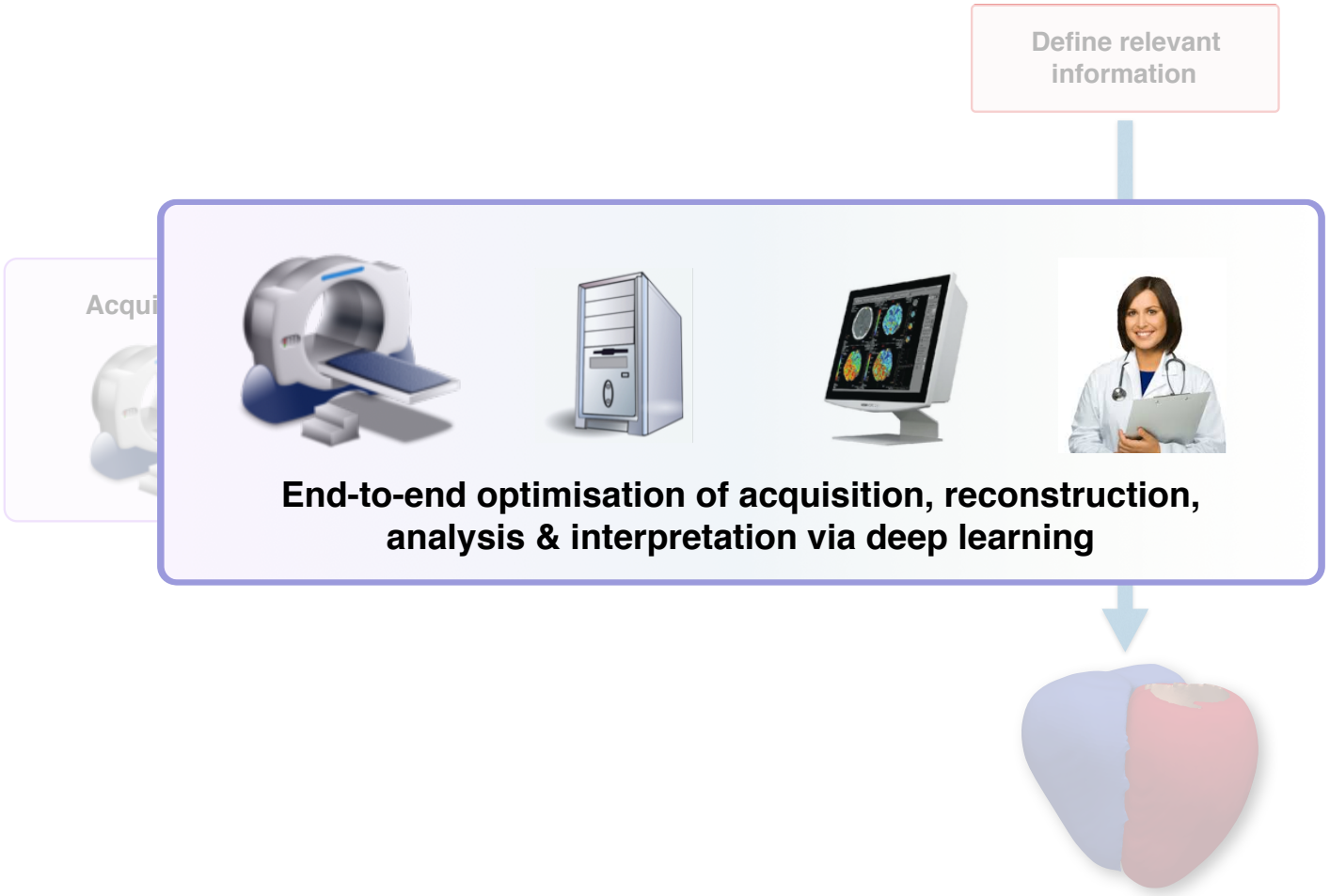
Optimisation of imaging pipeline with respect to clinically useful information



Current state-of-the-art



Future: End-to-end optimisation of entire imaging pipeline via deep learning



Future: End-to-end optimisation of entire imaging pipeline via deep learning

End-to-end optimisation of acquisition, reconstruction, analysis & interpretation via deep learning

biobank uk

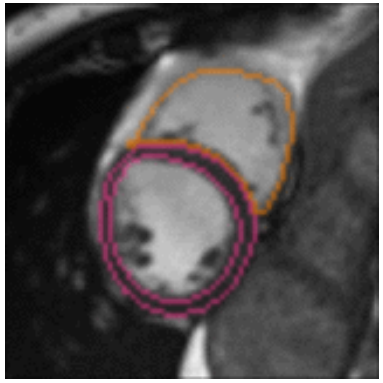
Dementias Platform UK
Medical Research Council

Big data (population data)

Multi-modal data

Sneak preview: Can we recover LV and RV directly from k-space?

Ground truth



J. Schlemper et al.
MICCAI 2018

Acknowledgements

Ben Glocker
Bernhard Kainz
Wenjia Bai
Matthew Sinclair
Giacomo Tarroni
Ozan Oktay
Martin Rajchl
Aaron M. Lee
Nay Aung
Elena Lukaschuk
Mihir M. Sanghvi
Filip Zemrak

Kenneth Fung
Jose Miguel Paiva
Valentina Carapella
Young Jin Kim
Hideaki Suzuki
Paul M. Matthews
Steffen E. Petersen
Stefan K. Piechnik
Stefan Neubauer
Enzo Ferrante
Steven McDonagh
Ghislain Vaillant

Jo Hajnal
Jose Caballero
Christian Ledig
Christian Baumgartner
Kostas Kamnitsas
Tong Tong
Wenzhe Shi
Martin Rajchl
Jo Schlemper
Carlo Biffi
Nick Pawlowski
Matthew Lee

This research has been conducted mainly using the UKBB Resource under Application Number 2946. The initial stage of the research was conducted using the UKBB Resource under Application Number 18545.

Acknowledgements

European Research Council