Imperial College London

Deep learning in medical imaging: Techniques for image reconstruction, super-resolution and segmentation

Daniel Rueckert

Biomedical Image Analysis Group

Department of Computing

Imperial College London, UK

Machine learning in medical imaging: There is a lot of hype

AI In Medicine: Rise Of The Machines (Forbes, 2017)

THE NEW YORKER

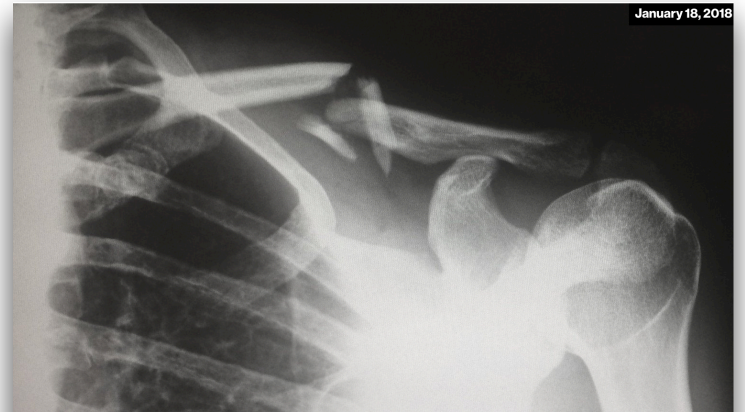
APRIL 3, 2017 ISSUE

A.I. VERSUS M.D.

What happens when diagnosis is automated?

By Siddhartha Mukherjee

MIT Technology Review



AI Is Continuing Its Assault on Radiologists

A new model can detect abnormalities in x-rays better than radiologists—in some parts of the body, anyway.

Machine learning in medical imaging:
There is a lot of hype

“They should stop training radiologists now.”
Geoffrey Hinton (godfather of deep learning) in 2017

"To the question, will AI replace radiologists, I say the answer is no…”

“… but radiologists who do AI will replace radiologists who don’t."
Curtis Langlotz in 2017


RSNA News

Machine Learning Plays Central Role at RSNA 2017

BY MIKE BASSETT

November 1, 2017

in Share | Tweet | f Share | @ Email

Langlotz

Machine Learning (ML) and the role it will play in the future of radiology will be central to a broad scope of programming at RSNA 2017.

Machine learning for medical imaging: Opportunities

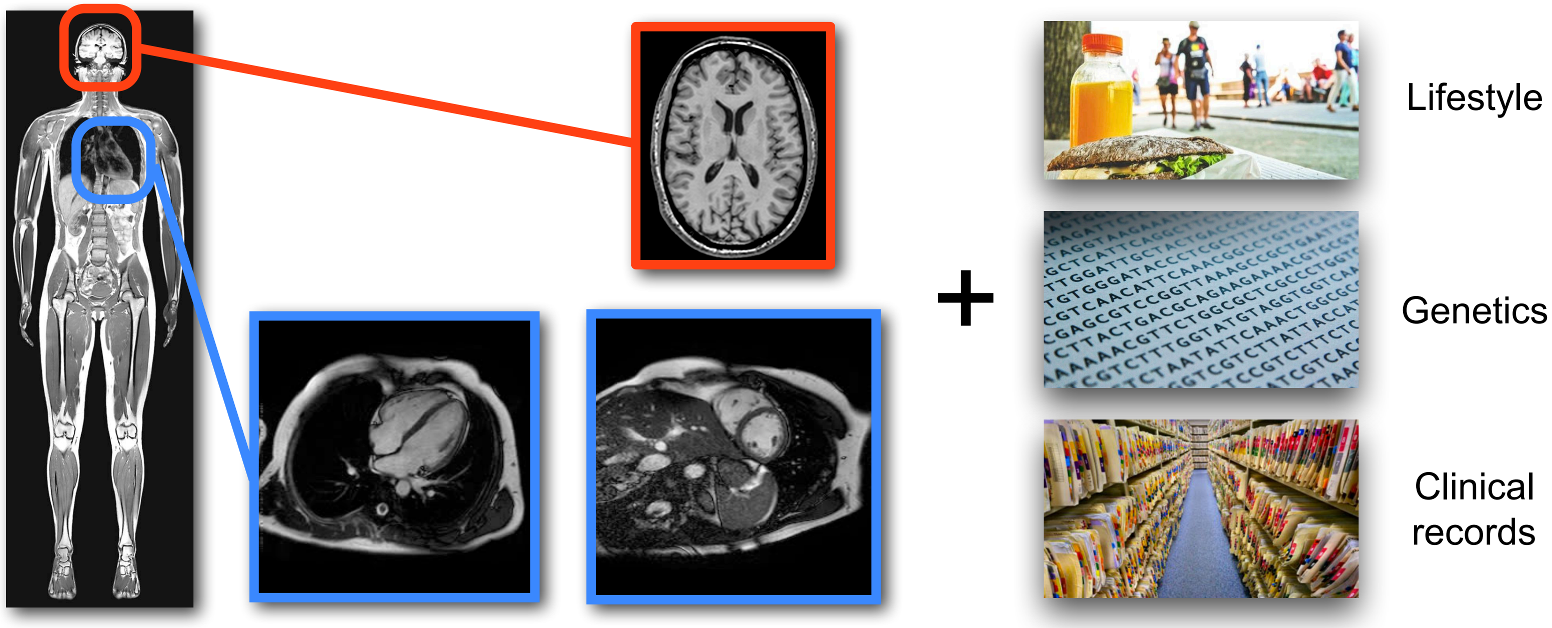
- Big data is slowly arriving in medical imaging

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

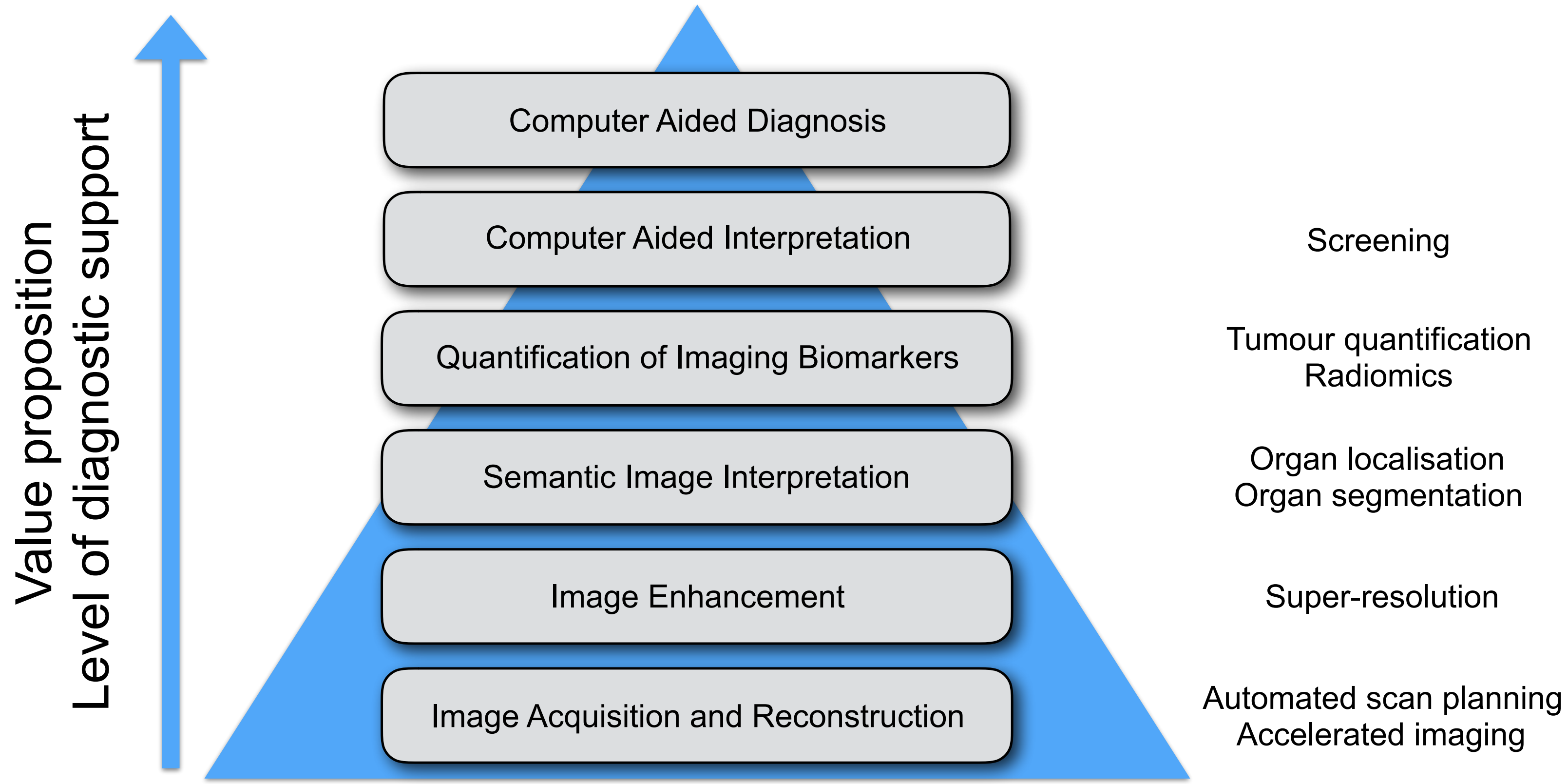
Machine learning for medical imaging: Opportunities

- Big data is slowly arriving in medical imaging

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



Machine learning for medical imaging: Opportunities



Practical challenges for ML in medical imaging

- Most ML approaches are supervised:
  Training data is key
- How to obtain training data?
- Training data is expensive:
  - manpower, cost, time
  - years of training and
- Traini ect:
  - training wrongly labelled

Crowdsourcing

Experts

If training data is not perfect how do we validate?

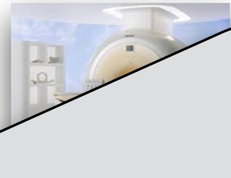
Practical challenges for ML in medical imaging

- ML-based solutions often degrade when deployed in clinical scenarios

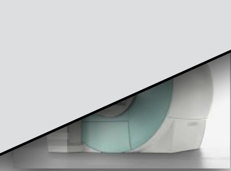
- This is caused by differences bet… training and test data, e.g.
  - different scanner hardw…
  - scanner protoco…
  - artefacts

- Manua… …otating new data for each test domain is not a feasible solution

Transfer learning problem



Philips



Siemens

GE

Overview

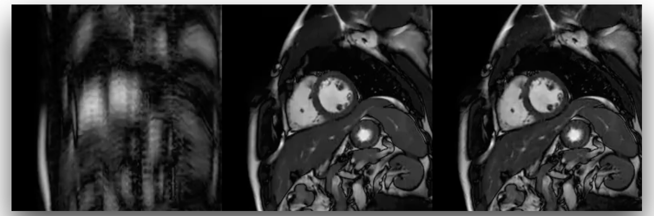


Image reconstruction

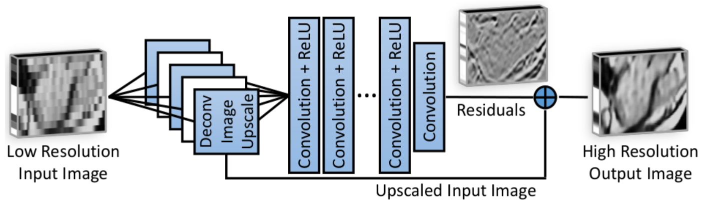


Image super-resolution

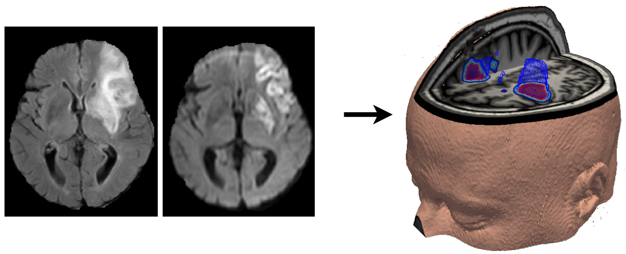
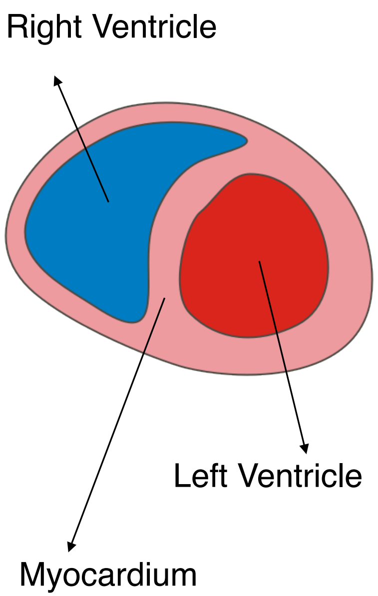


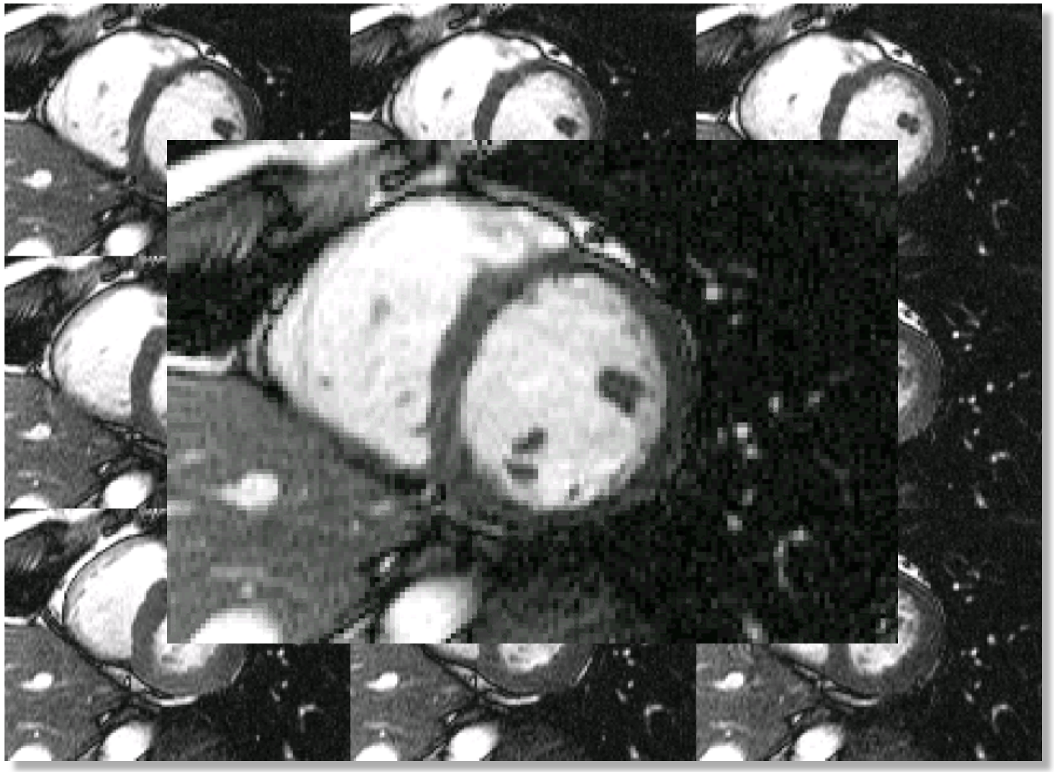
Image segmentation

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 sequenti
traversing sampling trajectories.

K-space

t = 0

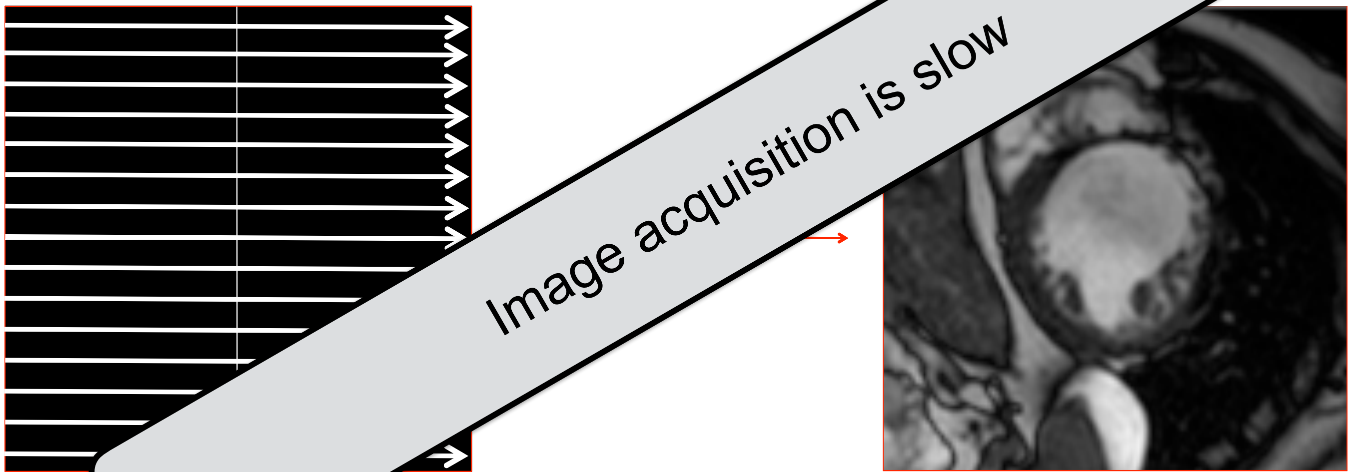
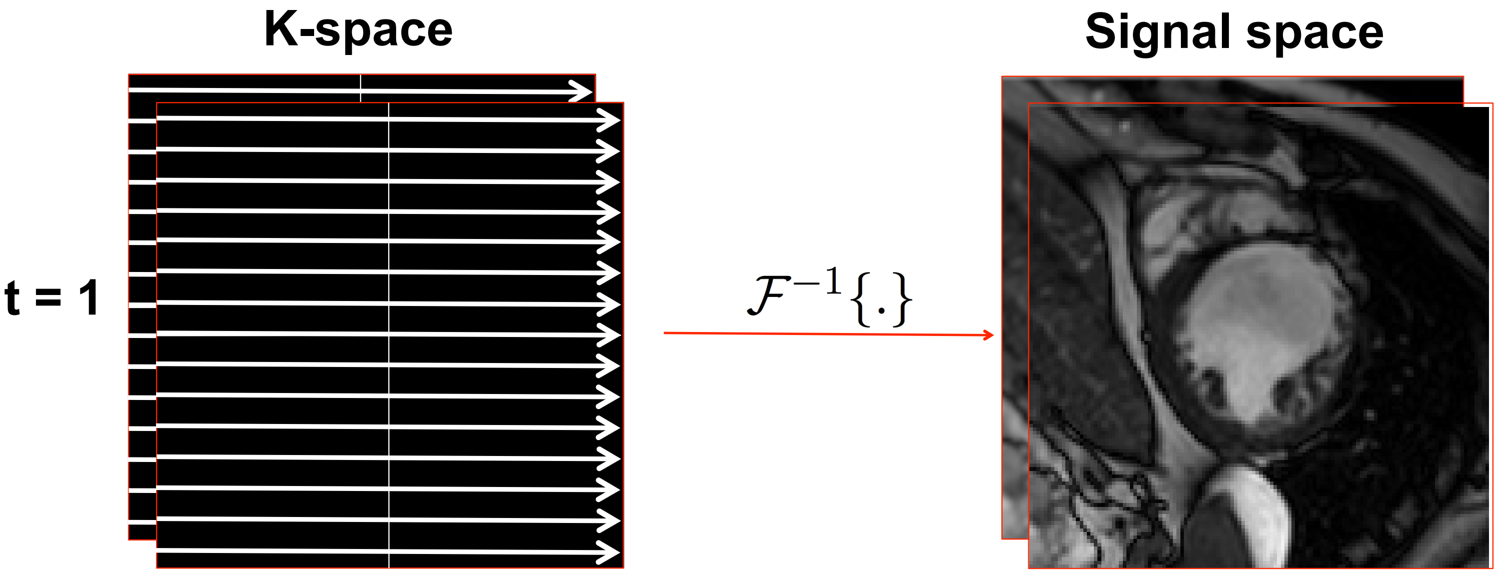


Image acquisition is slow

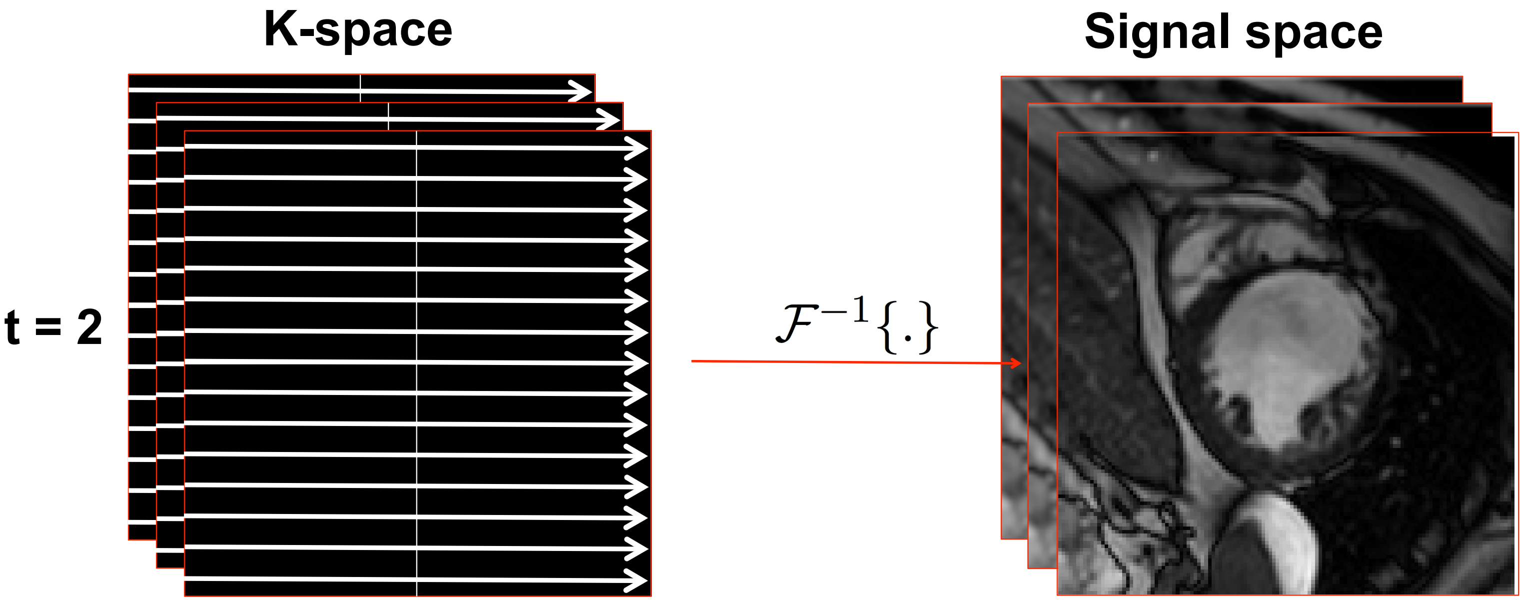
Cardiac MRI: Full acquisition is slow

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



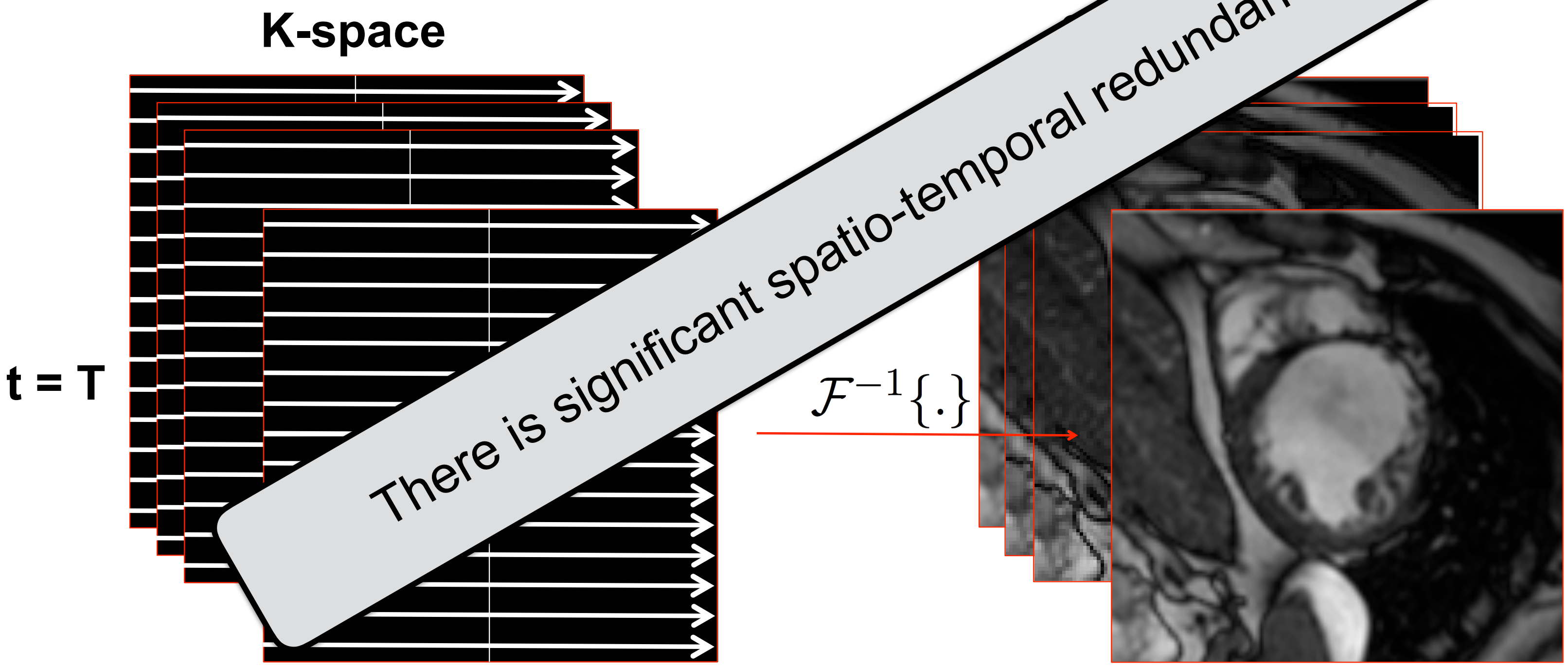
Cardiac MRI: Full acquisition is slow

- 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 sequenti
  traversing sampling trajectories.



K-space undersampling

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

K-space undersampling

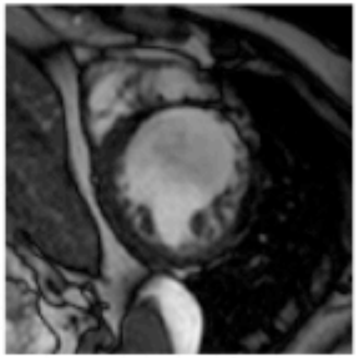
- Acquiring a fraction of k-space **<u>accelerates</u>** the process but introduces **<u>aliasing</u>** in signal space.

K-space

Signal space

Full sampling (slow)

$\mathcal{F}^{-1}\{.\}$



$\mathbf{x}_f$

learning-based reconstruction

e.g. compressed sensing

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
  - K. Hammernik et al., MRM 2017: Deep learning for knee MRI

Based on learnt priors

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 $\mathbf{F}_u = \mathbf{MF}$ 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 l0 or l1 norm)

Data fidelity term

- For CNN based reconstruction we formulate the problem as

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

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

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

Data consistency layer

- End-to-end training requires specification of forward and backward passes
- Forward pass:

$$f_L(\mathbf{x}, \mathbf{y}; \lambda) = \mathbf{F}^H \Lambda \mathbf{F} \mathbf{x} + \frac{\lambda}{1+\lambda} \mathbf{F}_u^H \mathbf{y}$$

- Backward pass:

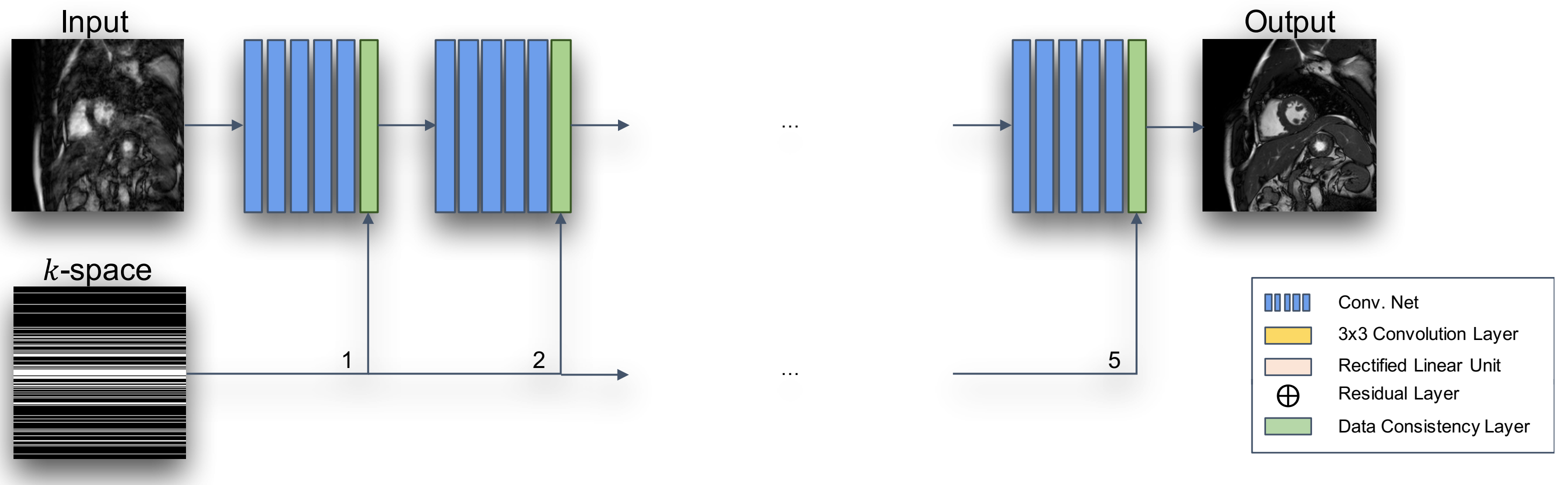
$$\frac{\partial f_L}{\partial \mathbf{x}^T} = \mathbf{F}^H \Lambda \mathbf{F}$$

Jacobian of the DC layer with respect to the layer input $\mathbf{x}$

$$\left[\frac{\partial f_{\mathrm{dc}}(\mathbf{s}, \mathbf{s}_0; \lambda)}{\partial \lambda}\right]_j = \begin{cases} 0 & \text{if } j \notin \Omega \\ \dfrac{\mathbf{s}_0(j) - \mathbf{s}_{\mathrm{cnn}}(j)}{(1+\lambda)^2} & \text{if } j \in \Omega \end{cases}$$

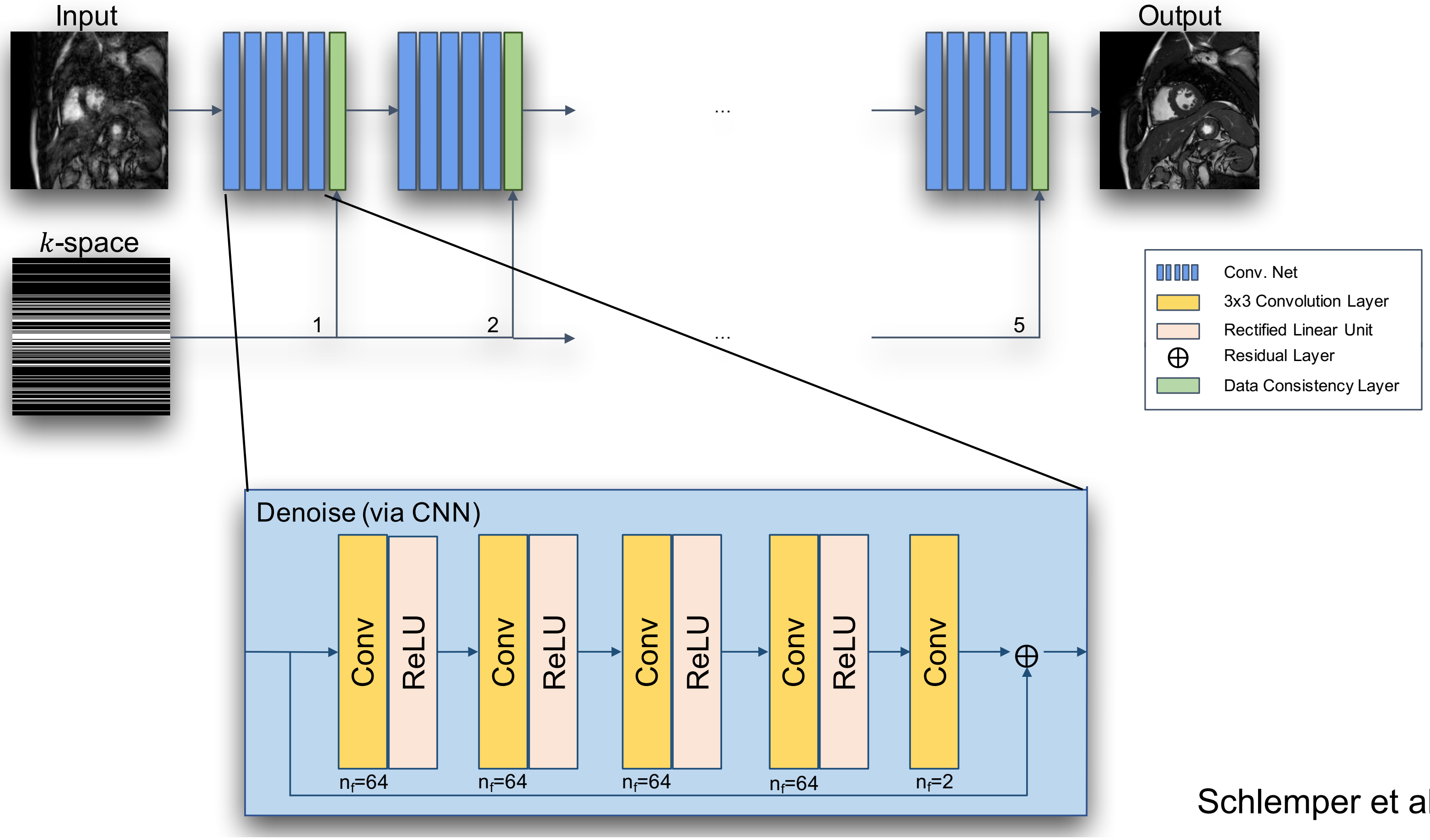
If made trainable

Deep Cascade of CNNs for MRI Reconstruction



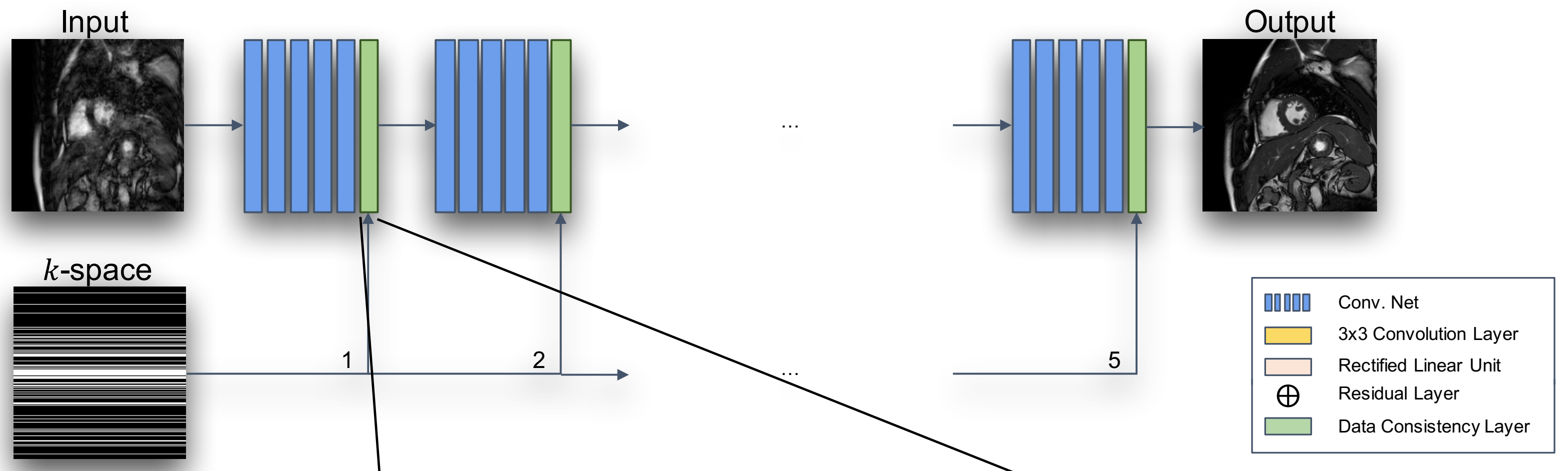
Schlemper et al. IEEE TMI 2017

Deep Cascade of CNNs for MRI Reconstruction



Schlemper et al. IEEE TMI 2017

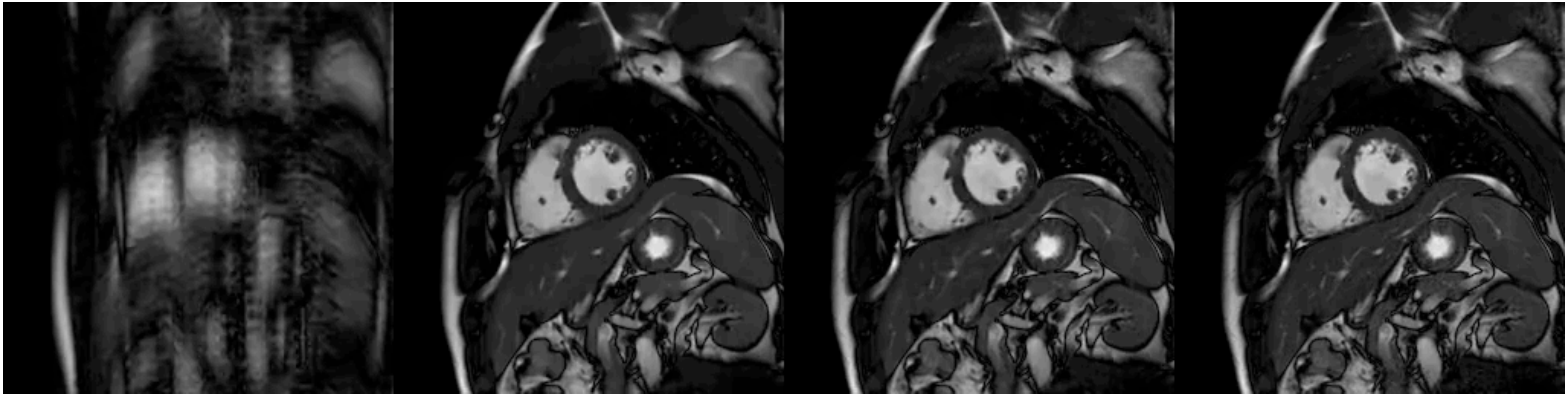
Deep Cascade of CNNs for MRI Reconstruction



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

Schlemper et al. IEEE TMI 2017

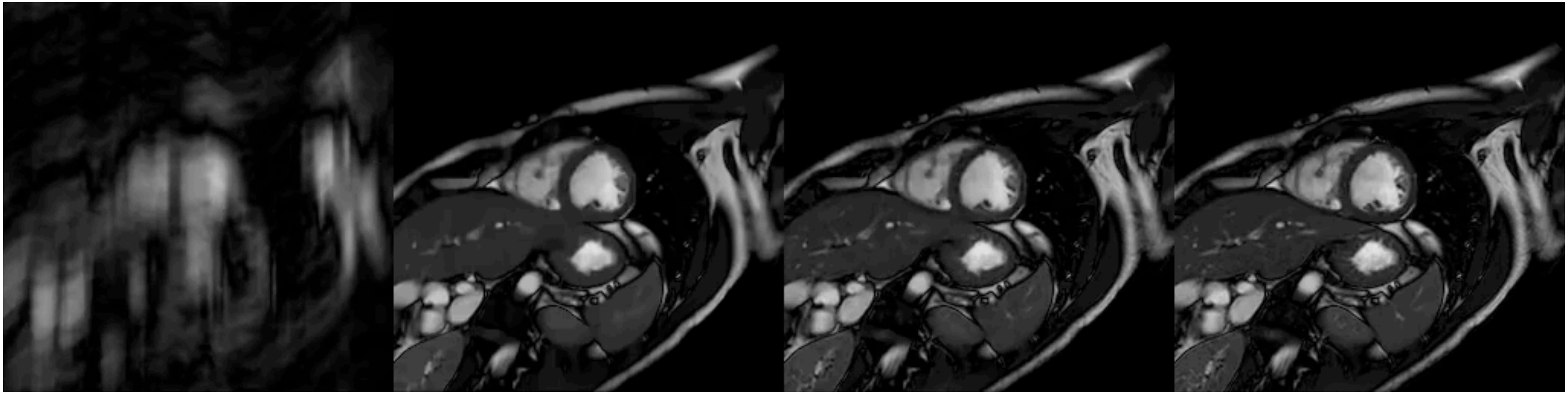
Magnitude reconstruction (6-fold)



(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

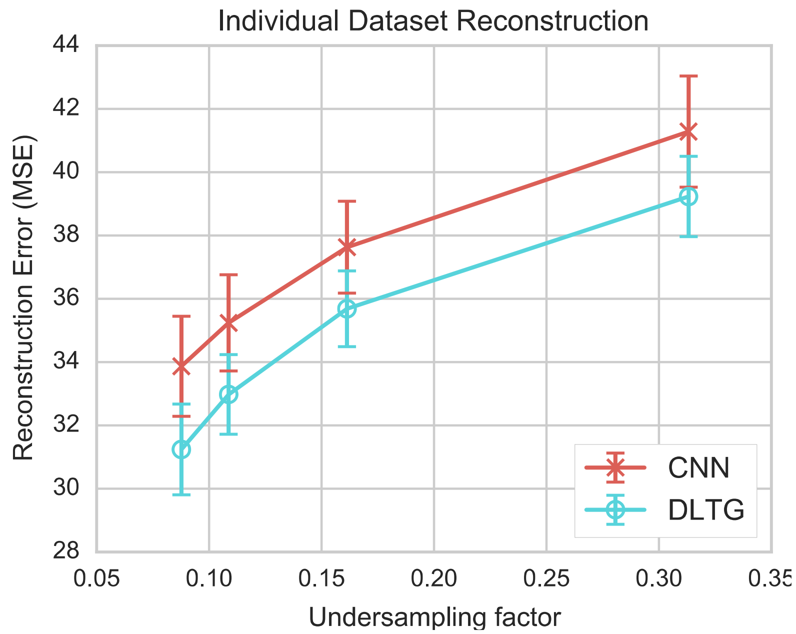
- Test error across 10 subjects:

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

Overview

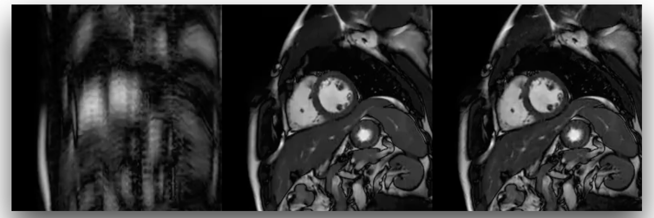


Image reconstruction

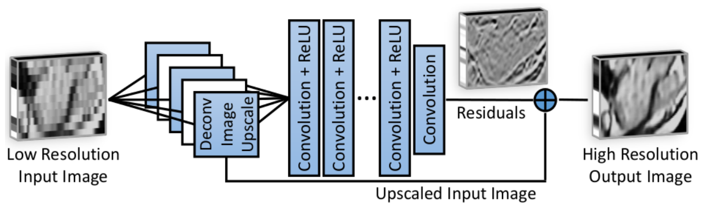


Image super-resolution

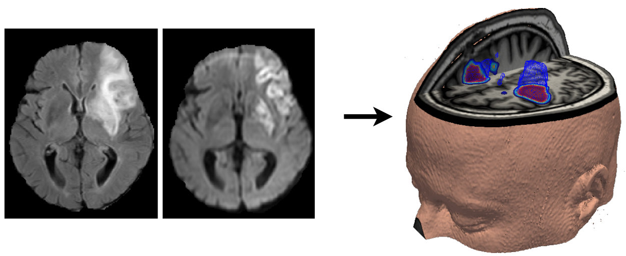
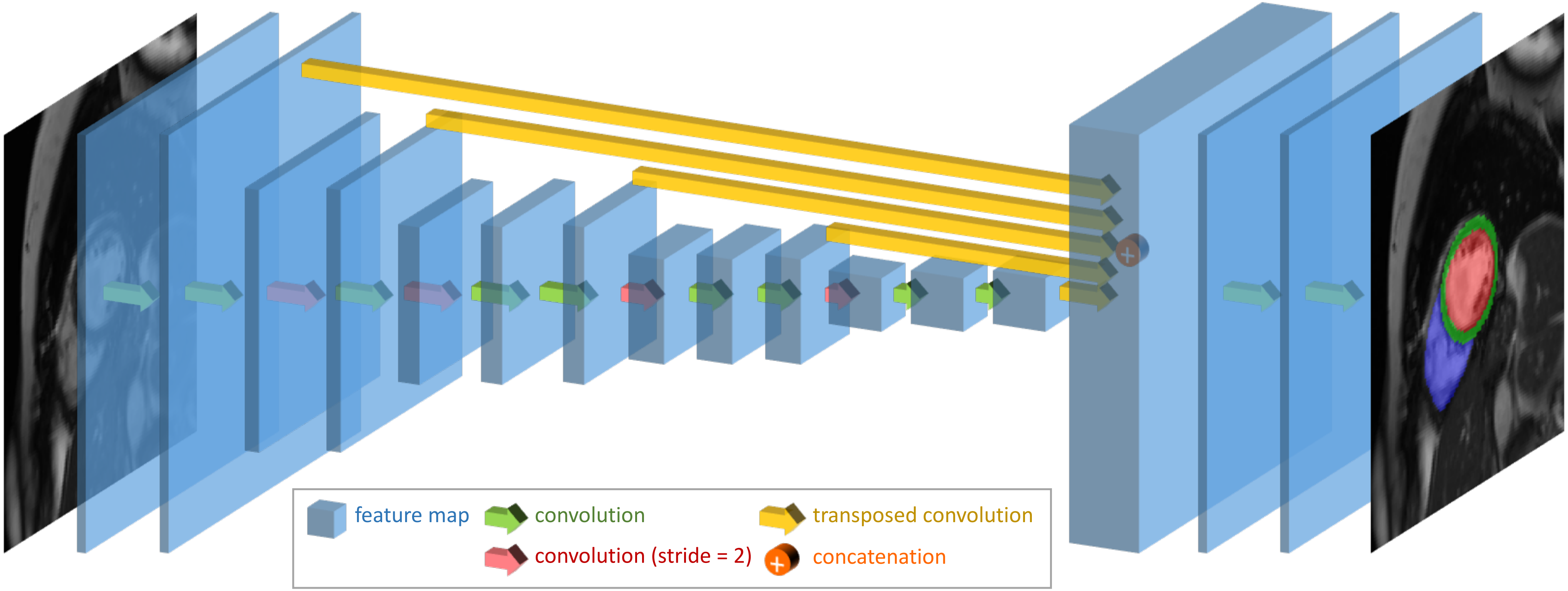


Image segmentation

Convolutional Neural Networks for Medical Image Segmentation



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

Image segmentation as a machine learning problem

- Manual annotations of **<u>4,872 subjects</u>** (QMUL/Oxford) with **<u>93,128 pixelwise annotated 2D images</u>** 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

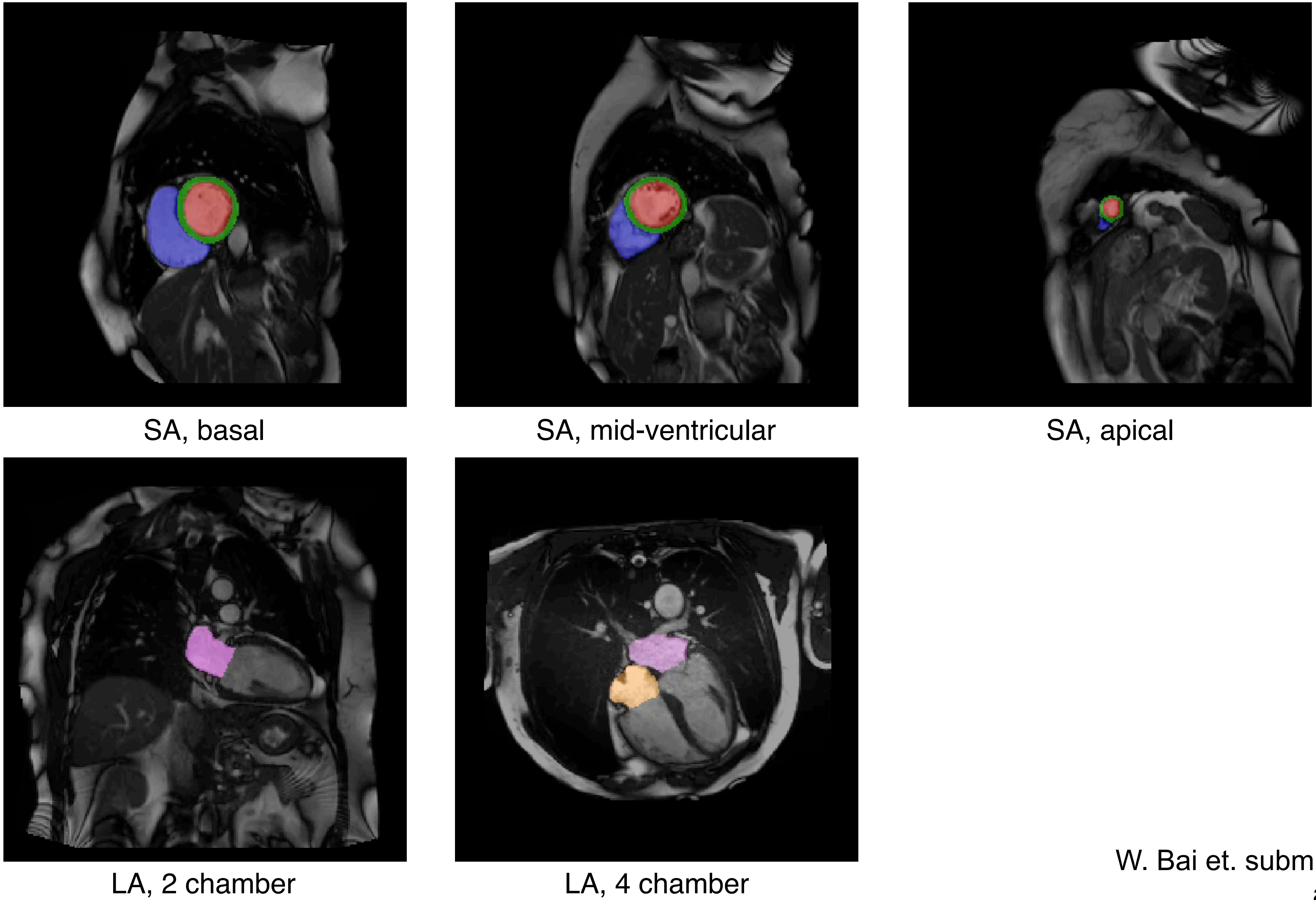
RESEARCH Open Access

CrossMark

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. submitted to JCMR, 2018
arXiv:1710.09289v3



W. Bai et. submitted to JCMR, 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 [illegible]
LVEDV (mL)	$6.1_{\pm 5.3}$	$6.1_{\pm 4.4}$	8.8[illegible]	[illegible]
LVESV (mL)	$5.3_{\pm 4.9}$	$4.1_{\pm 4.2}$	[illegible]	[illegible]
LVM (gram)	$6.9_{\pm 5.5}$	4.2[illegible]	[illegible]	[illegible]$_{\pm 4.8}$
RVEDV (mL)	$8.5_{\pm 7.1}$	[illegible]	[illegible]	$8.7_{\pm 5.8}$
RVESV (mL)	$7.2_{\pm 6.8}$	[illegible]	[illegible]$_{\pm 5.5}$	$11.7_{\pm 6.9}$

[illegible]nce

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

Automated

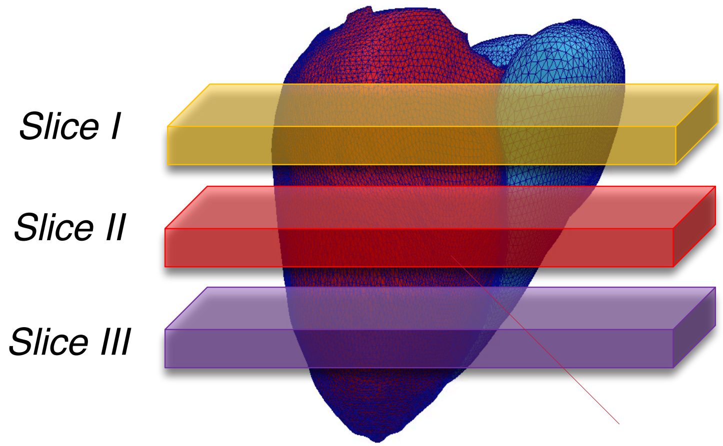
Manual

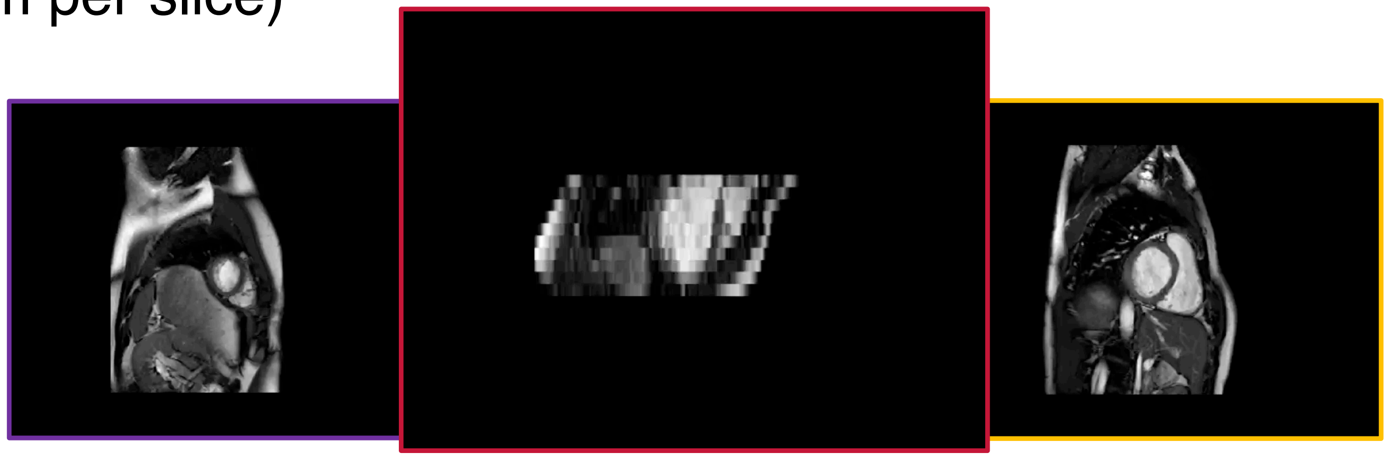
Computer performs as well as different expert observers

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

But: Cardiac imaging is still challenging

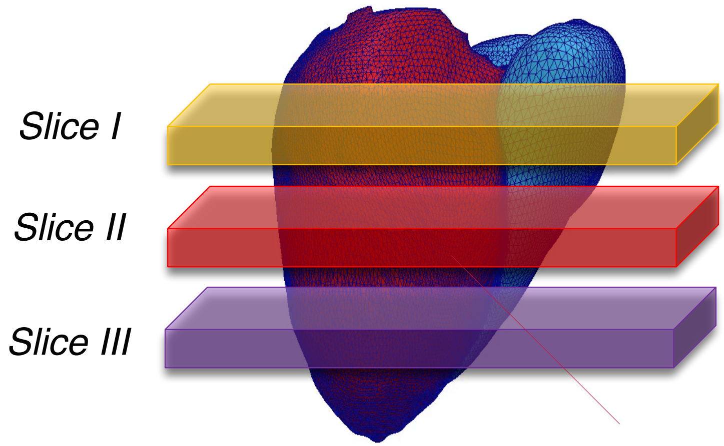
- Acquisition of cardiac MRI typically consists of 2D multi-slice data due to
  - constraints on SNR
  - breath-hold time
  - total acquisition time
- This leads to thick slice data (thickness 8-10 mm per slice)

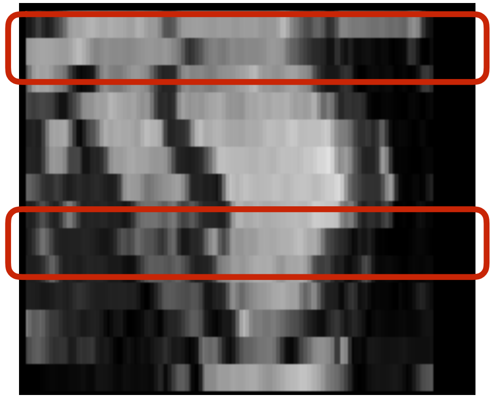




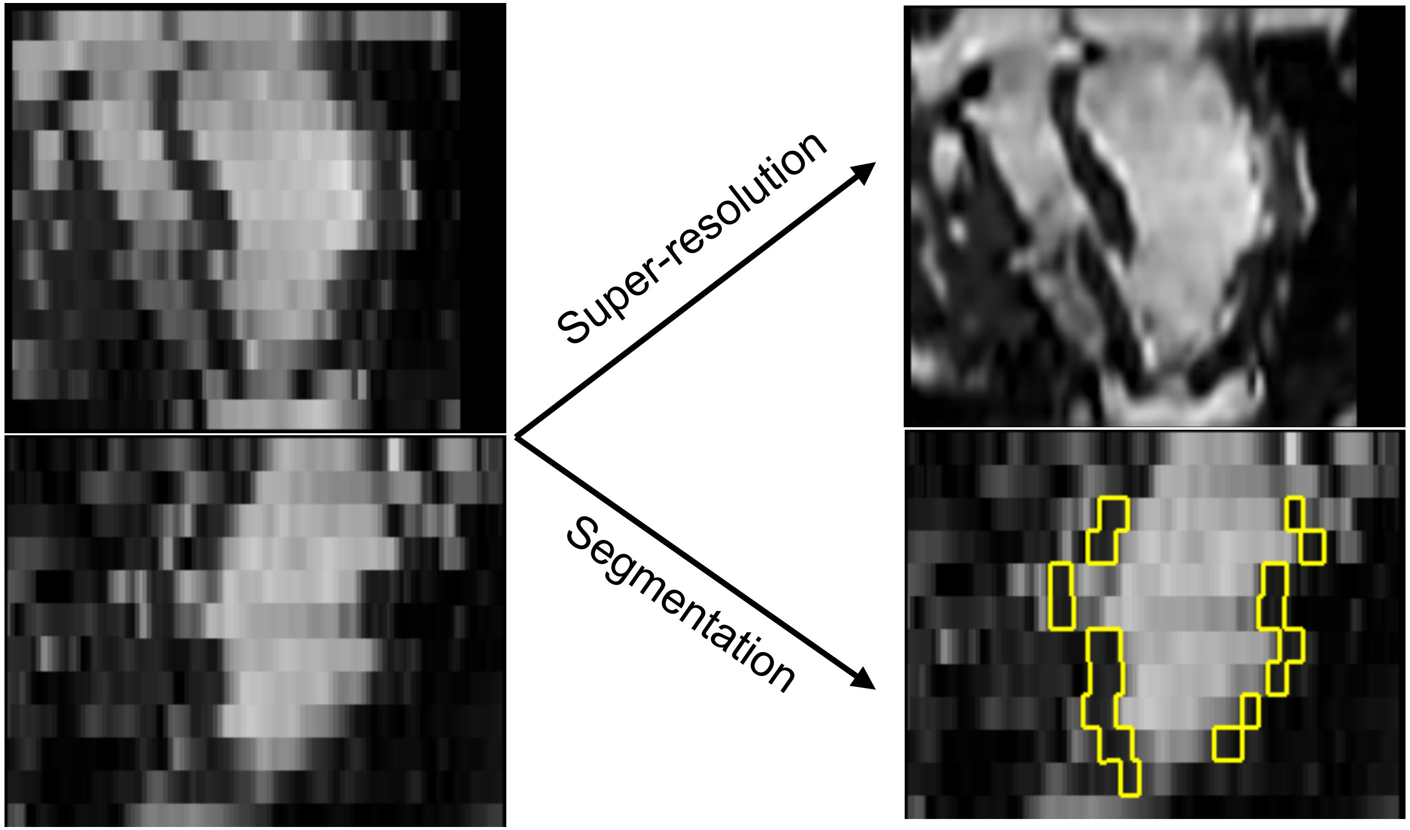
But: Cardiac imaging is still challenging

- 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

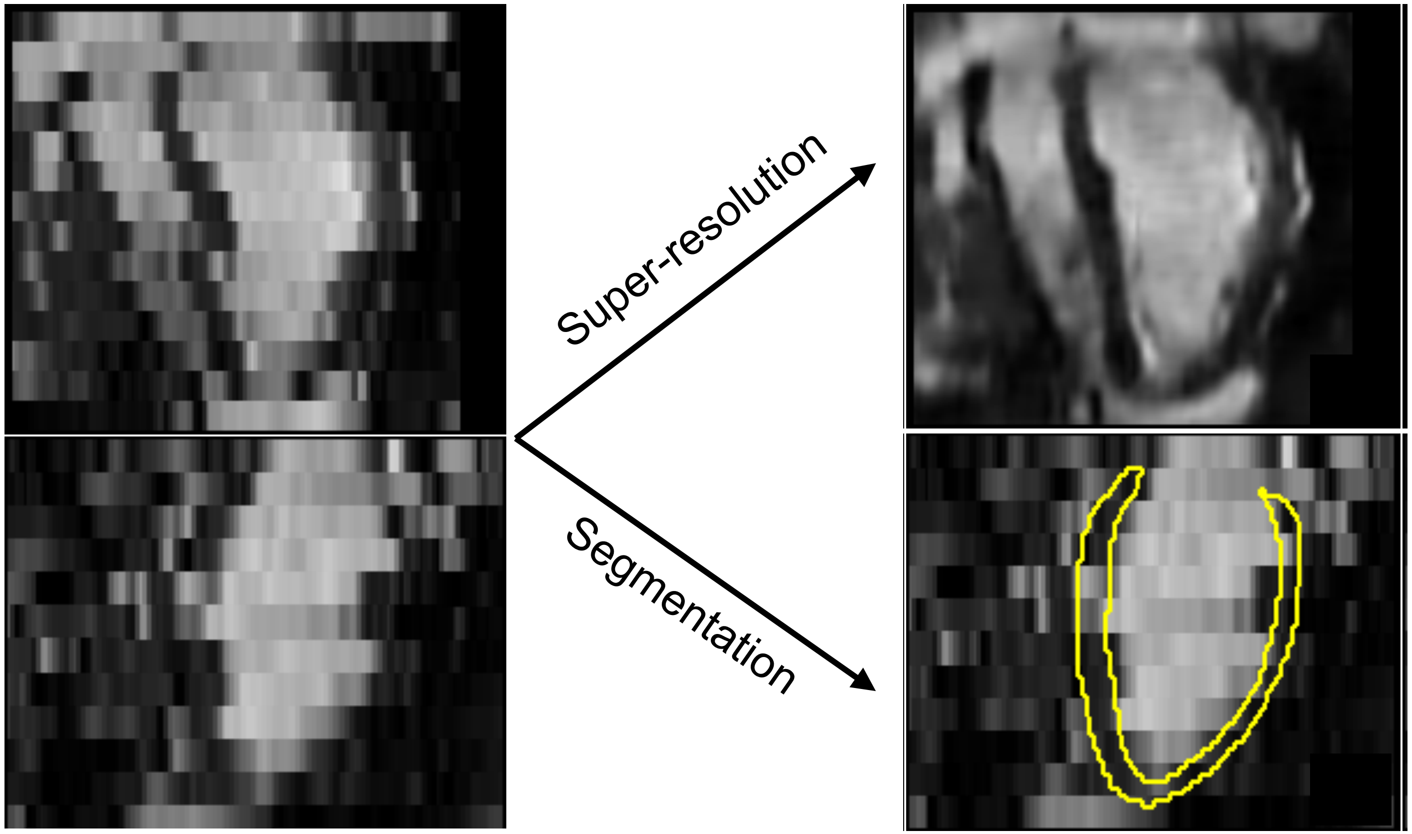




Conventional CNNs: Problem



Conventional CNNs: What we want



Conventional CNNs: No explicit use of prior knowledge

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

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

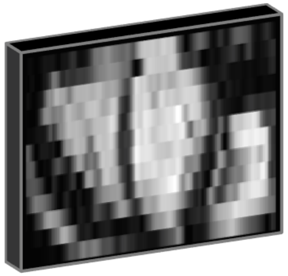
- Standard loss for **super-resolution**: L2 or L1 loss

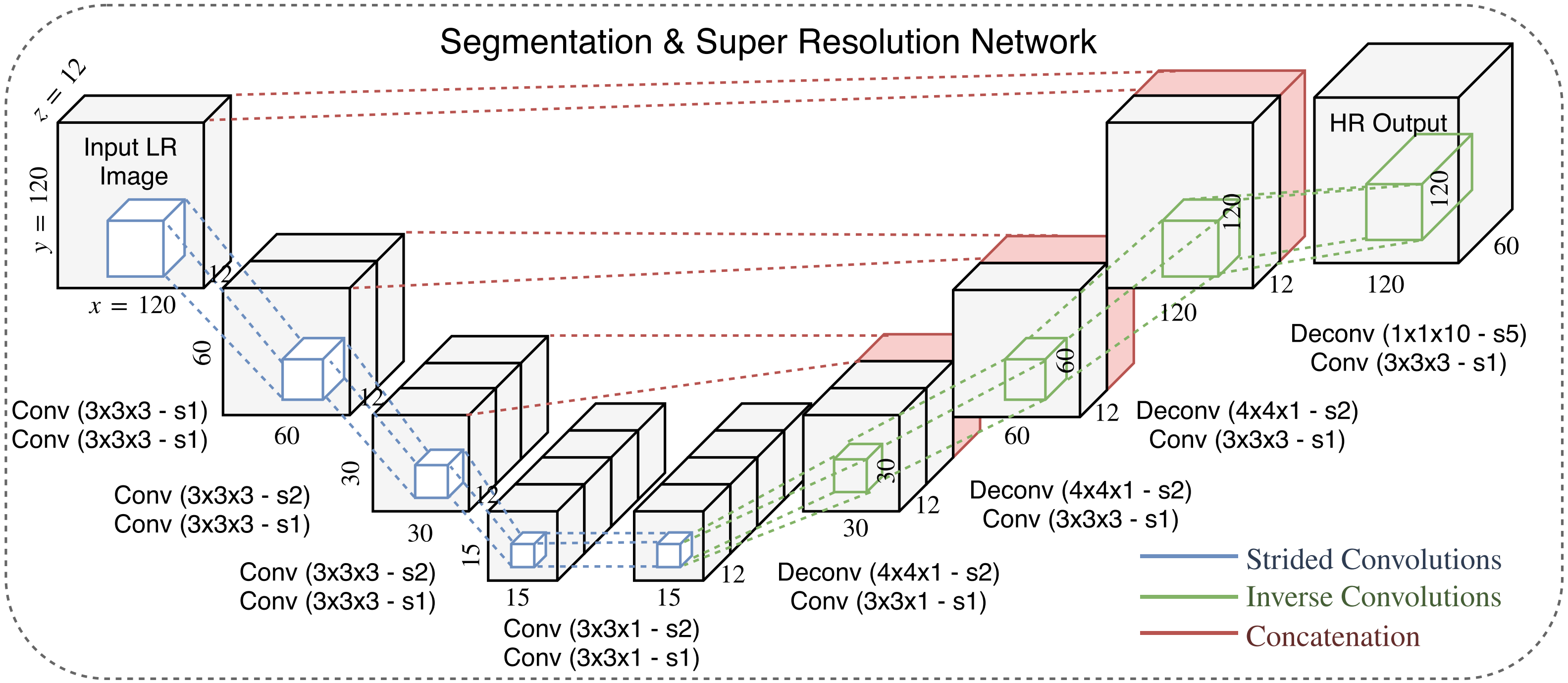
$$\sum_{i \in \mathcal{S}} \| \Phi(\boldsymbol{x}_i, \theta_r) - \boldsymbol{y}_i \|^2$$

Anatomically constrained CNNs

Low-resolution input

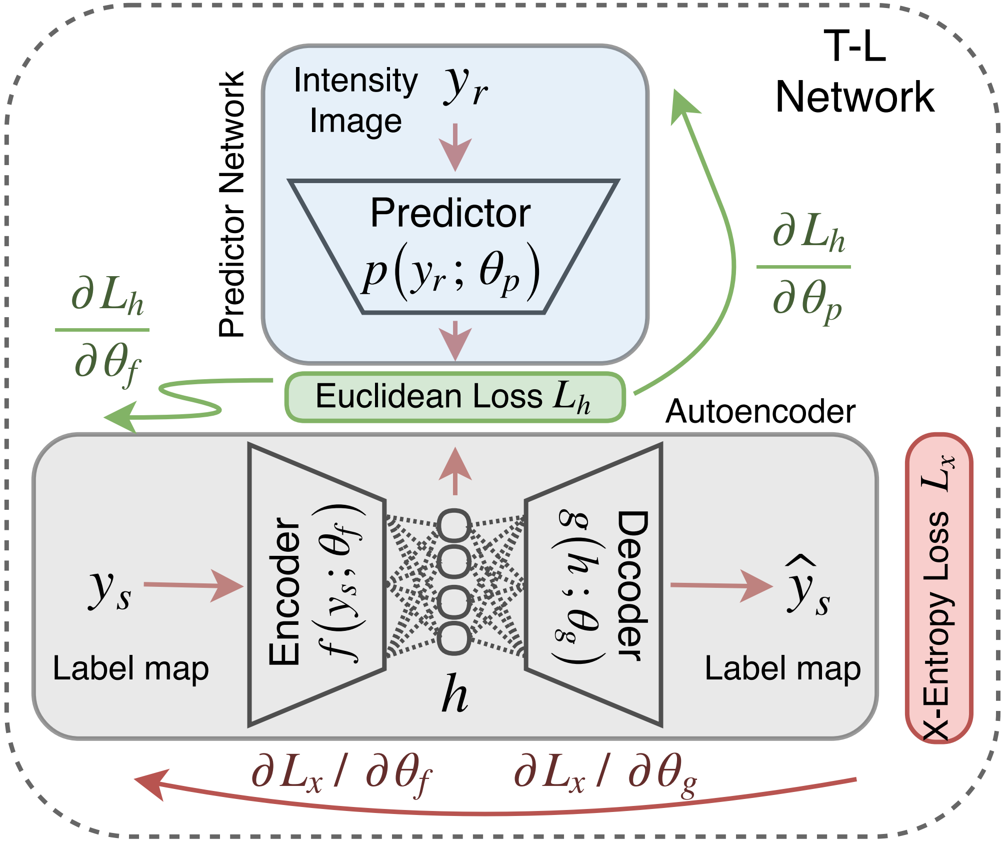
High-resolution output





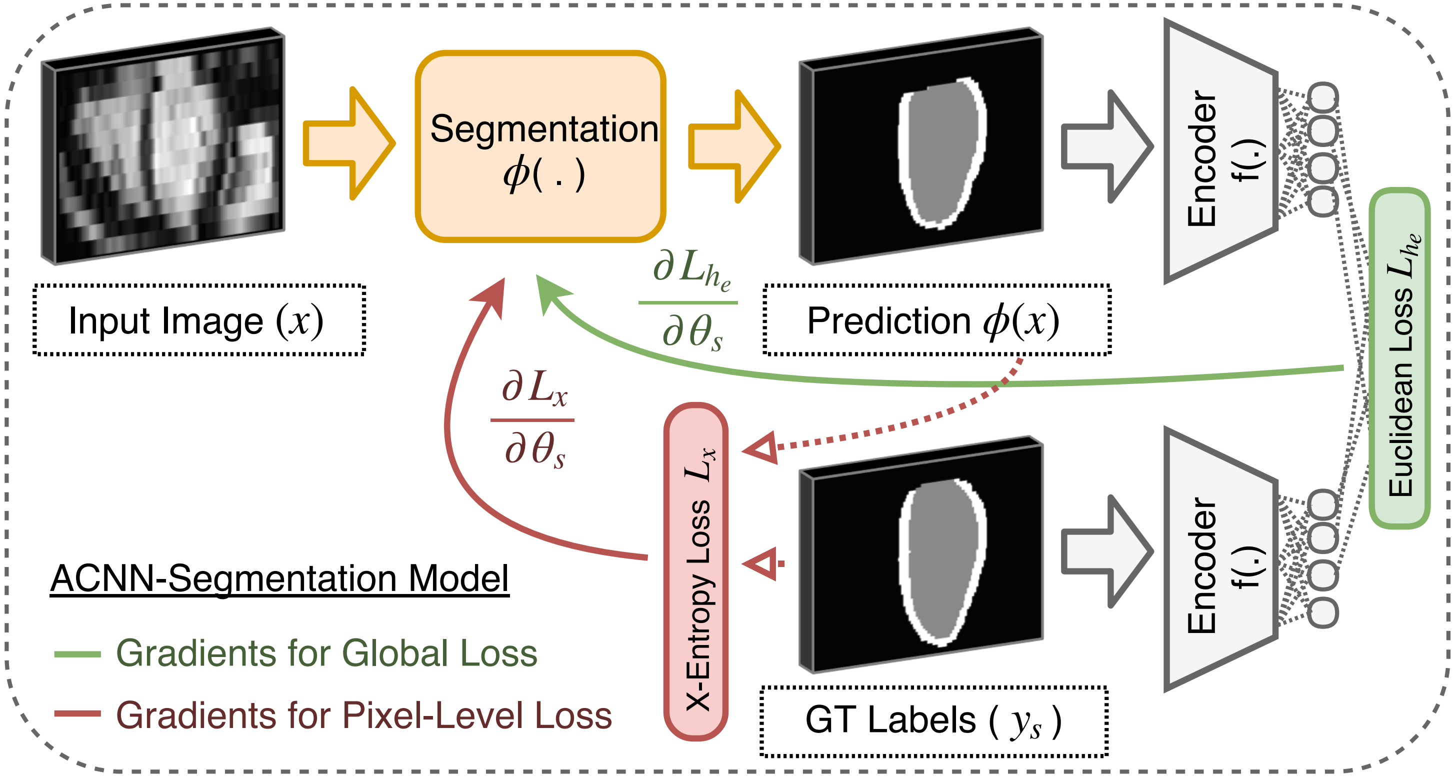
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: T-L networks for representing priors



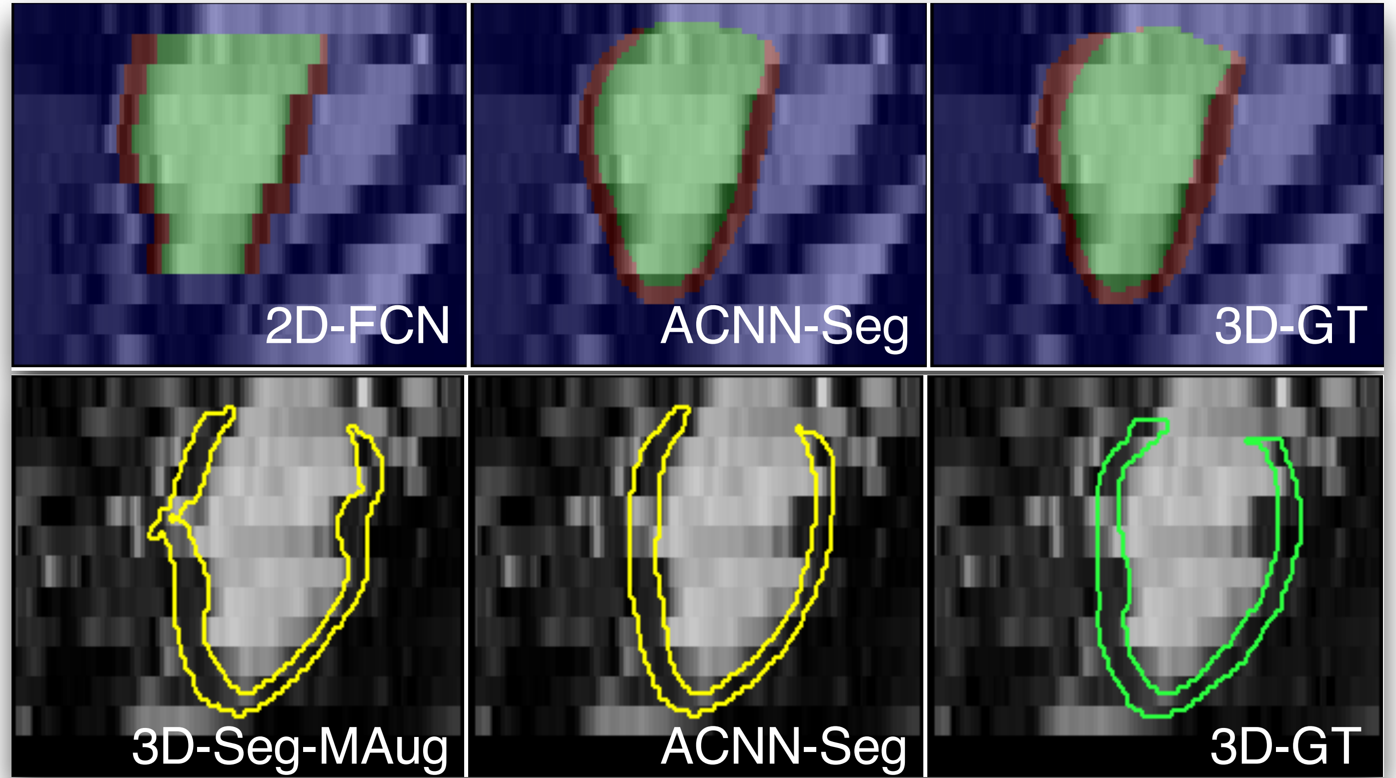
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Segmentation framework



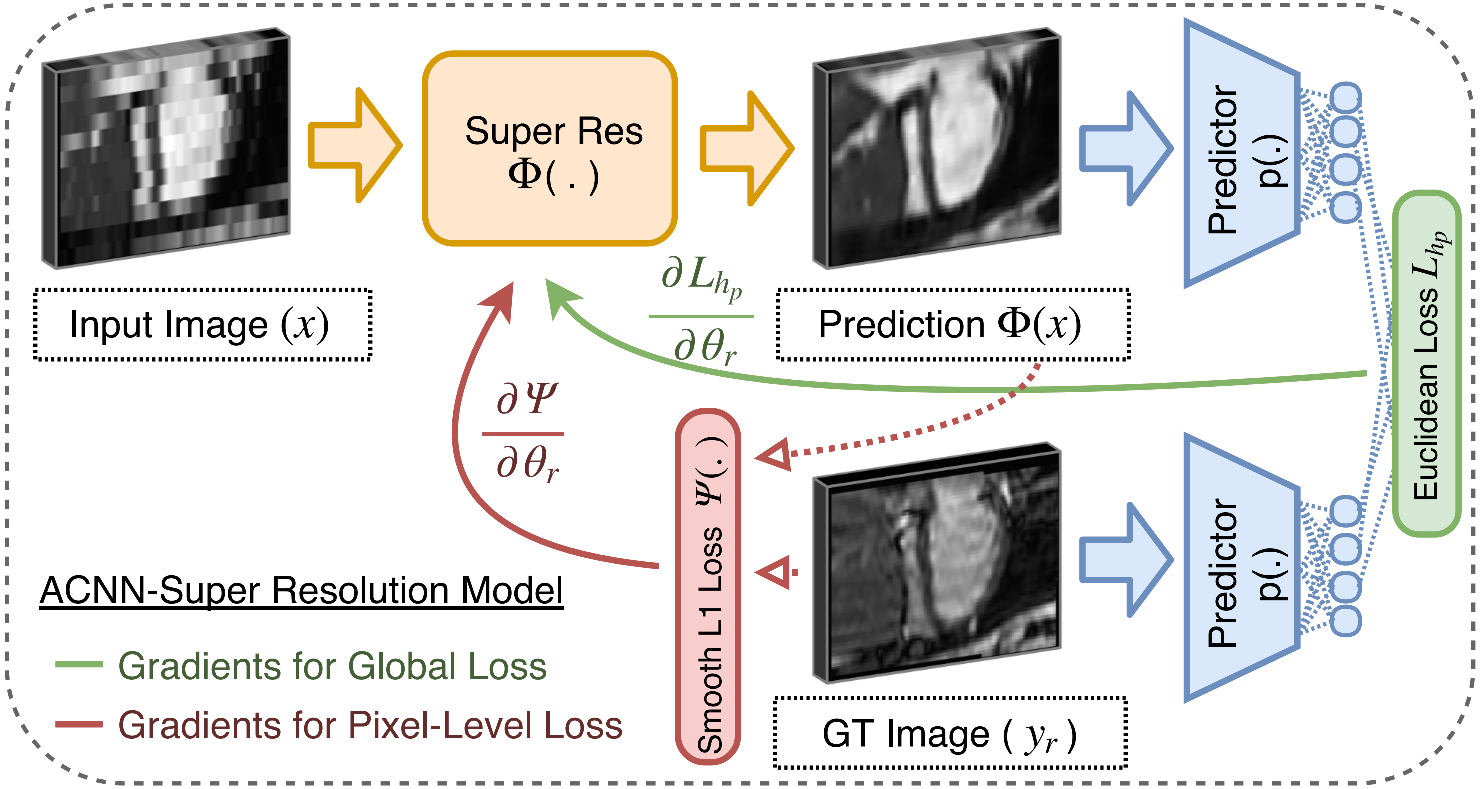
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Segmentation results



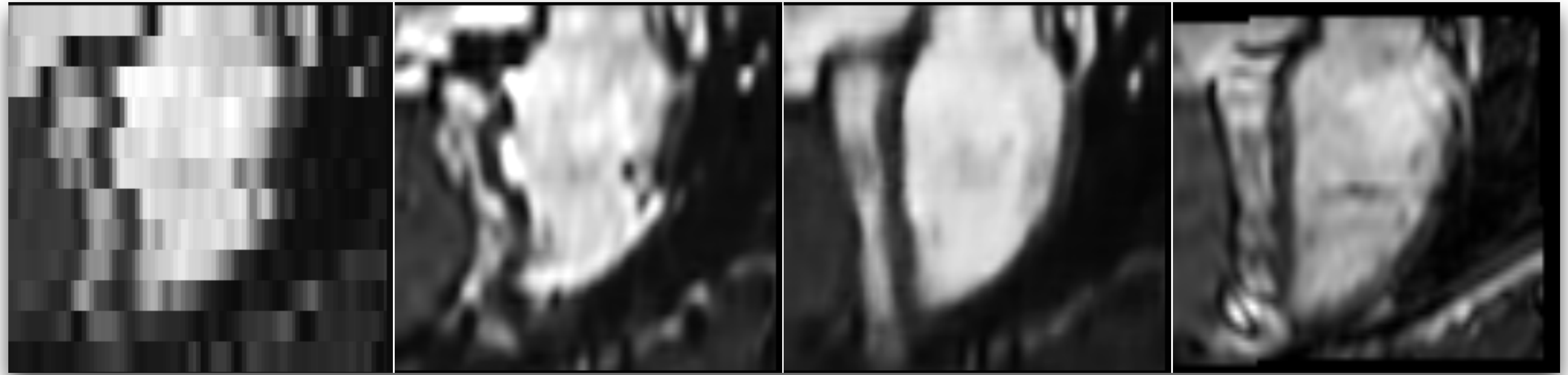
O. Oktay et al. IEEE TMI 2017

Anatomically constrained CNN: Super-resolution framework



O. Oktay et al. IEEE TMI 2017

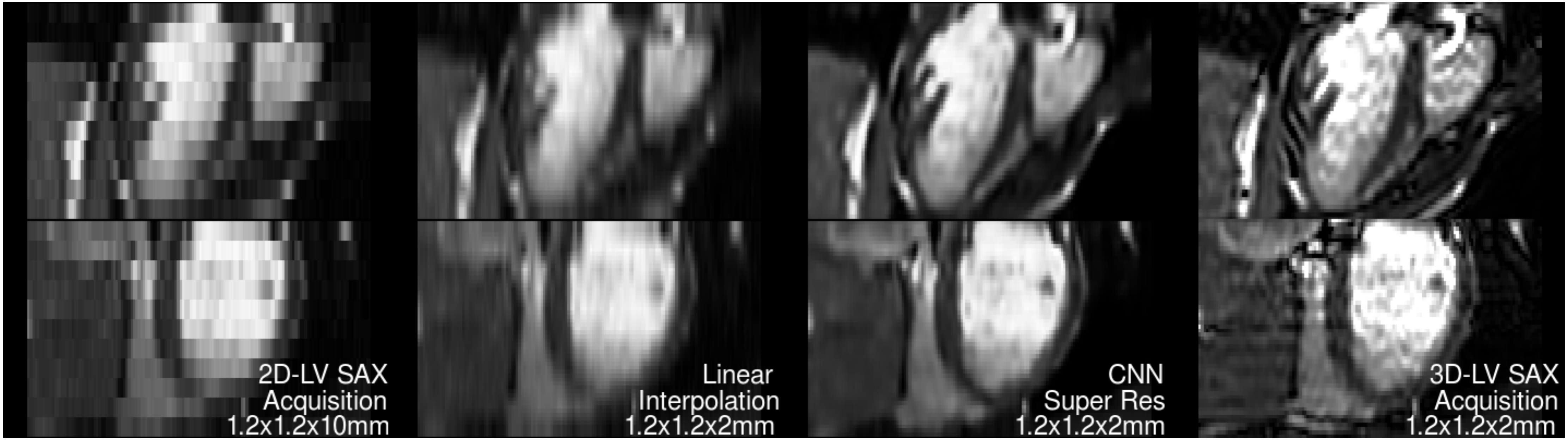
Anatomically constrained CNN: Super-resolution results



Original LR image | Baseline SR approach | Anatomically constrained SR model | Ground-truth HR image

O. Oktay et al. IEEE TMI 2017

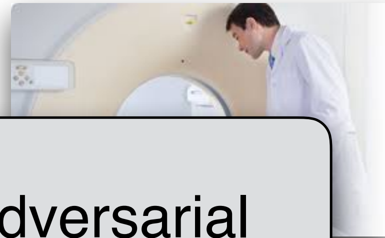
Anatomically constrained CNN: Super-resolution results



O. Oktay et al. IEEE TMI 2017

Challenges for medical image segmentation: Deployment in the clinic

- ML-based segmentation often degrades when deployed in clinical scenarios
- Thi
  trai
  var
  - s
  - s
- Ma
  test domain is not a feasible solution



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?

Source (S)

Domain: $D_S = \{\mathcal{X}_S, P(X_S)\}$

Task: $T_S = \{\mathcal{Y}_S, f'_S : \mathcal{X}_S \mapsto \mathcal{Y}_S\}$

Given: (X_S, Y_S)

$X_S = \{x_{S1}, ..., x_{Sn}\}, x_{Si} \in \mathcal{X}_S$

$Y_S = \{y_{S1}, ..., y_{Sn}\}, y_{Si} \in \mathcal{Y}_S$

Learn: $f_S \approx f'_S$

$f_S(x) \approx P_S(y|x)$

Target (T)

Domain: $D_T = \{\mathcal{X}_T, P(X_T)\}$

Task: $T_T = \{\mathcal{Y}_T, f'$

Here: $\mathcal{Y}_T = \mathcal{Y}$

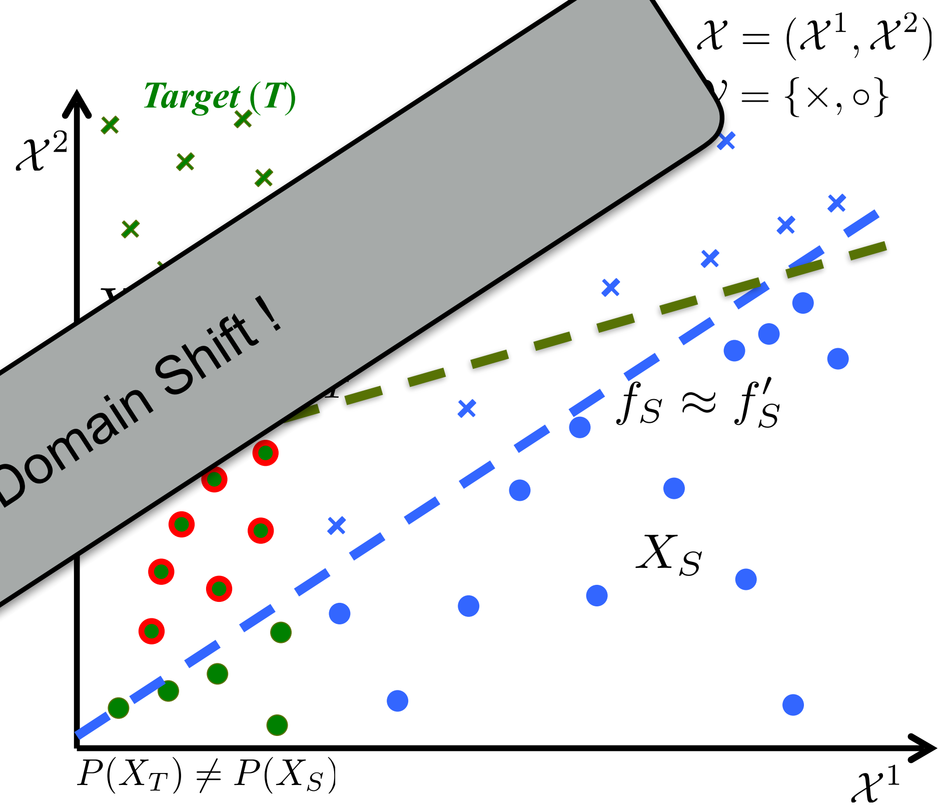
Domain Shift:

$P(X$

$P_T(y|x$ $|x)$

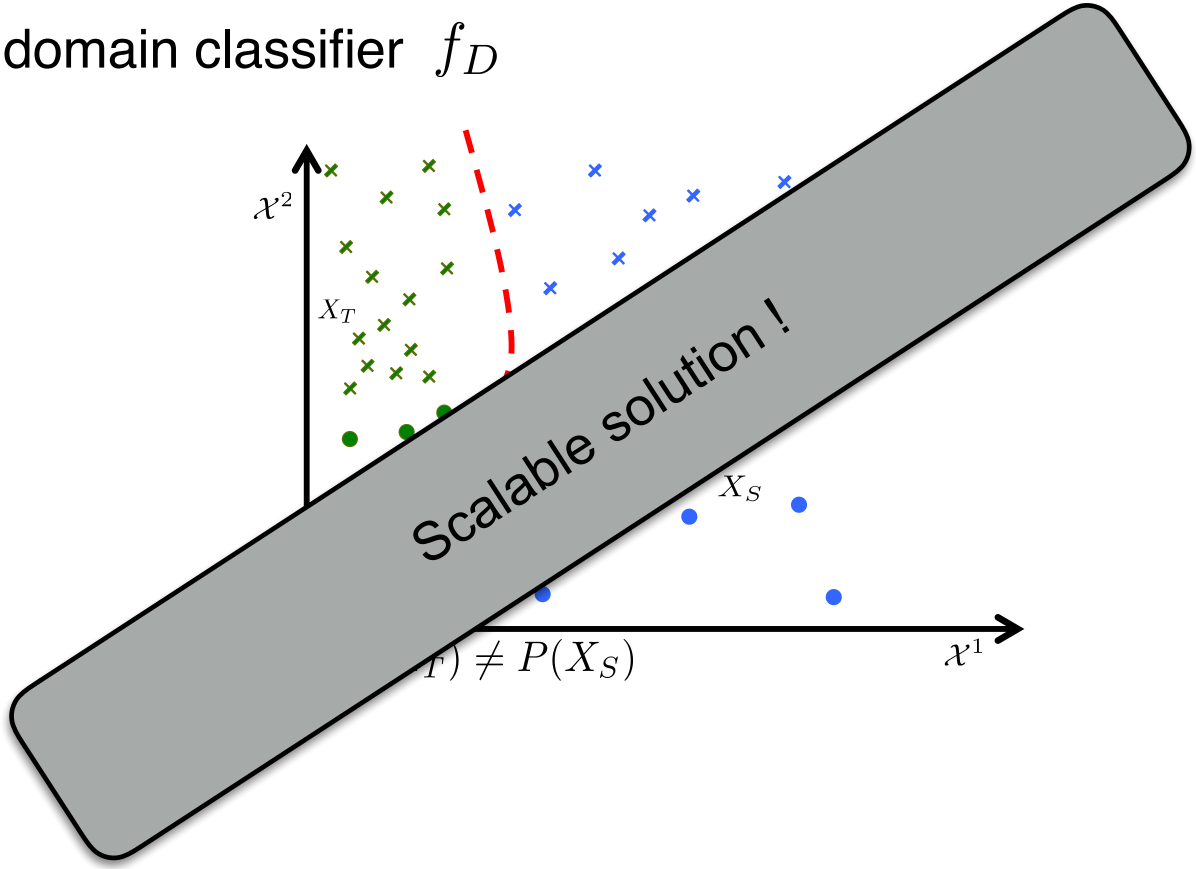
$f'_T \neq f'_S$

Domain Shift !

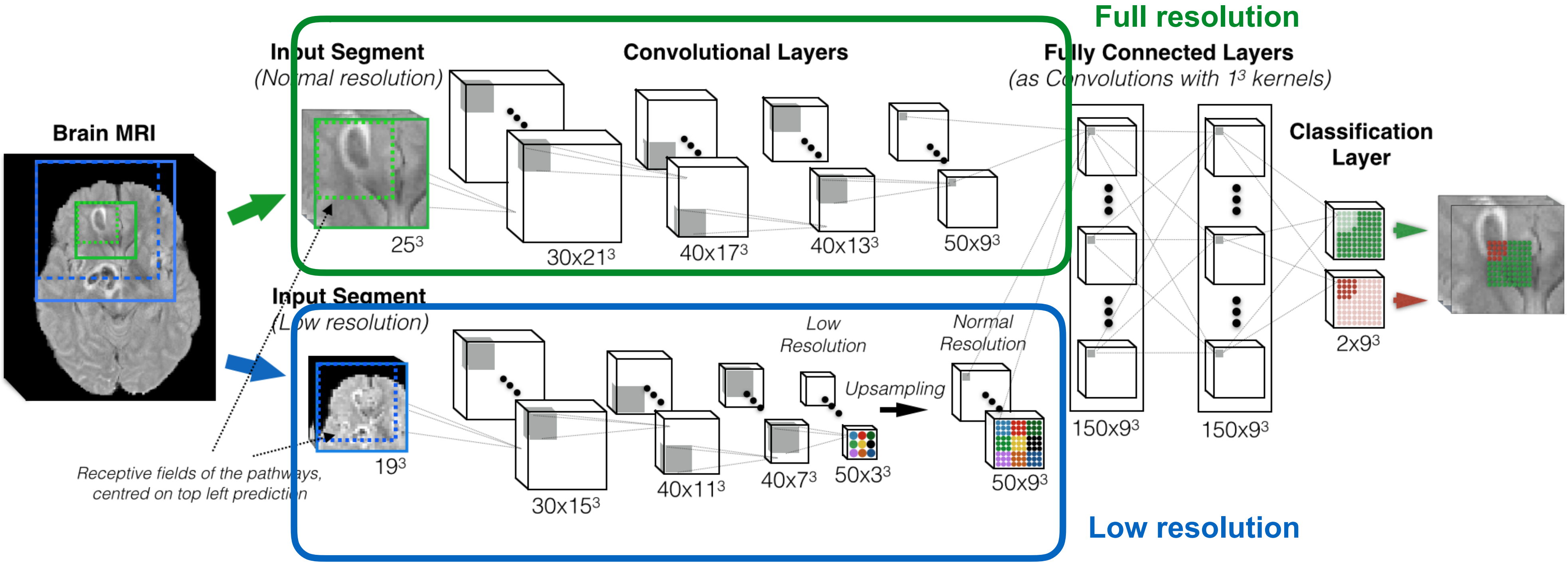


Solution: Unsupervised domain adaptation with adversarial networks

- Learn a domain classifier f_D

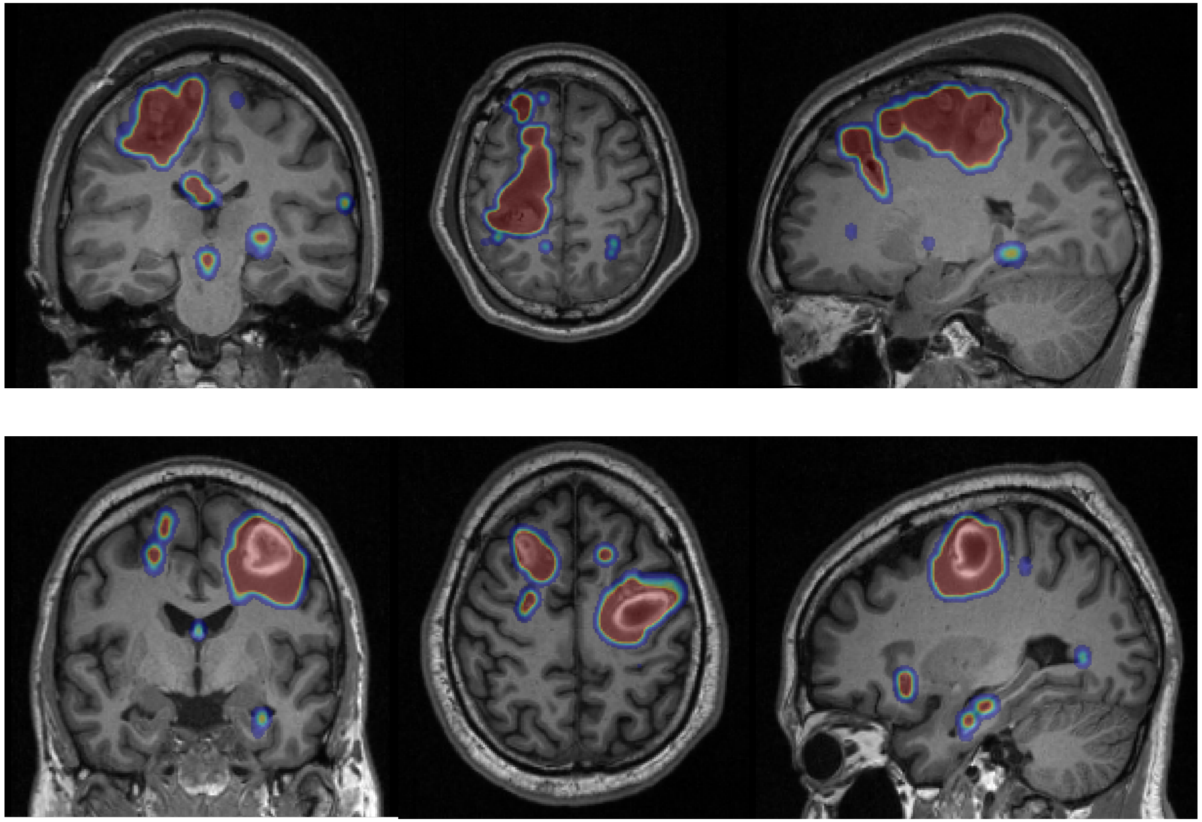


DeepMedic: Overview



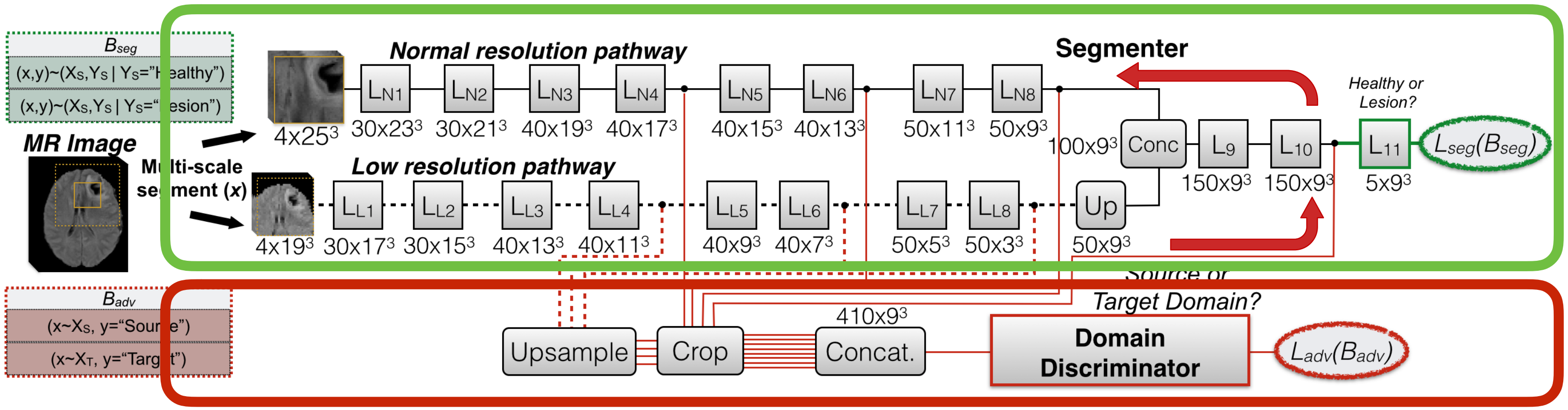
K. Kamnitsas et al. Medical Image Analysis, 2016

DeepMedic in Action

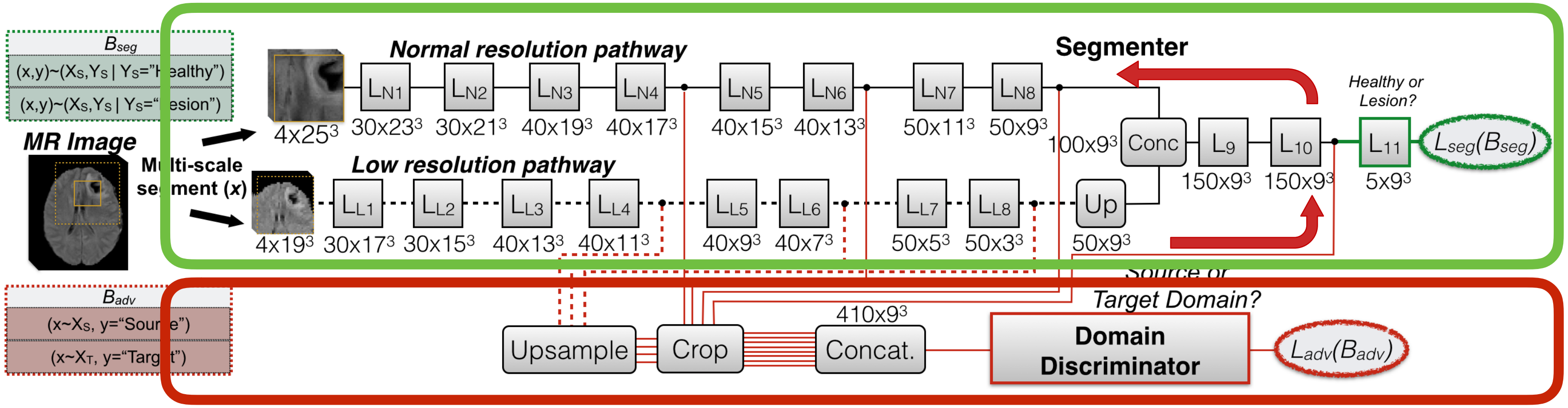


K. Kamnitsas et al. Medical Image Analysis, 2016

DeepMedic: Unsupervised domain adaptation with adversarial networks



DeepMedic: Unsupervised domain adaptation with adversarial networks



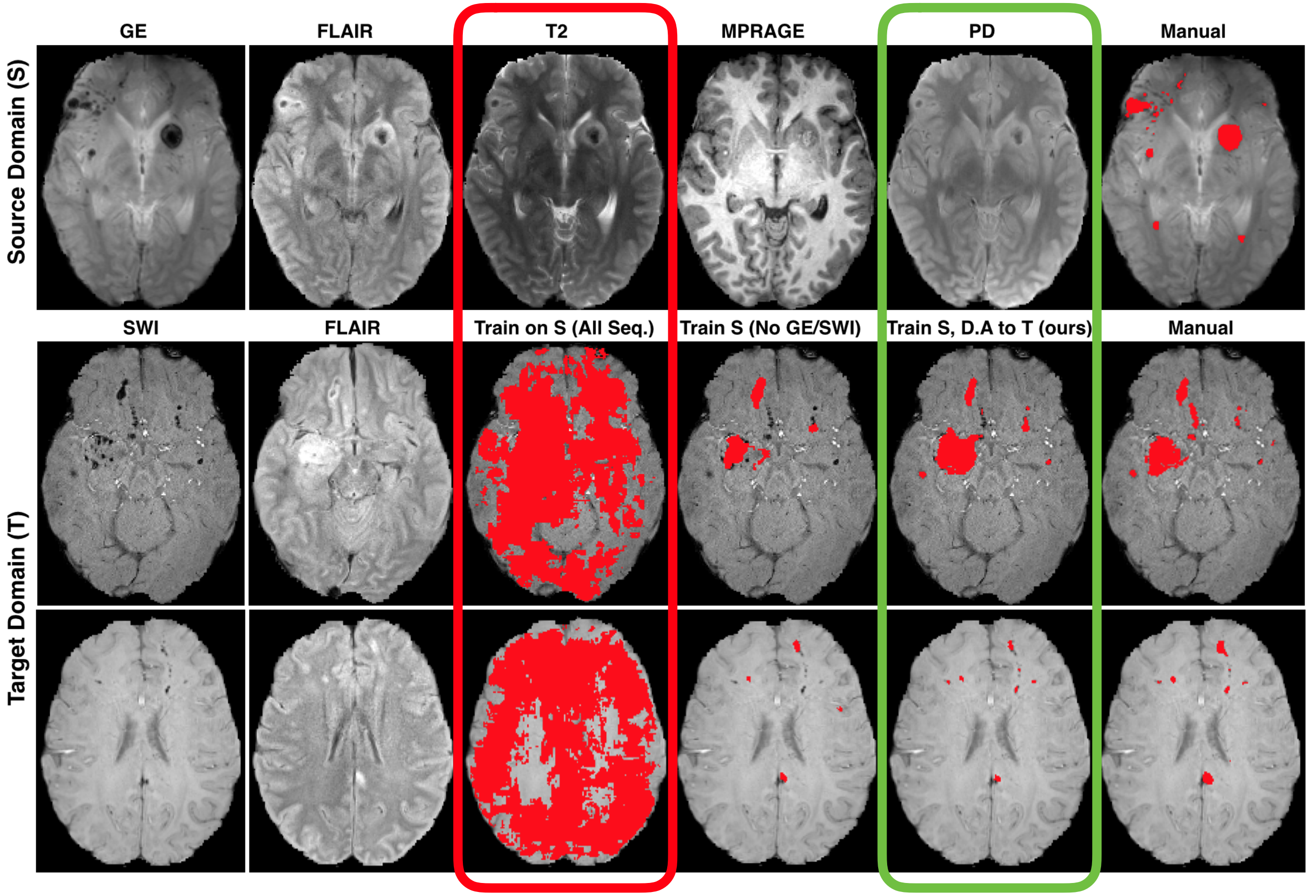
Segmenter: $\mathcal{L}_{seg} = -\frac{1}{m}\sum_{i=1}^{m}[f(x_i) = y_i]log(f(x_i)) \quad ,(x_i, y_i) \sim (X_S, Y_S)$

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

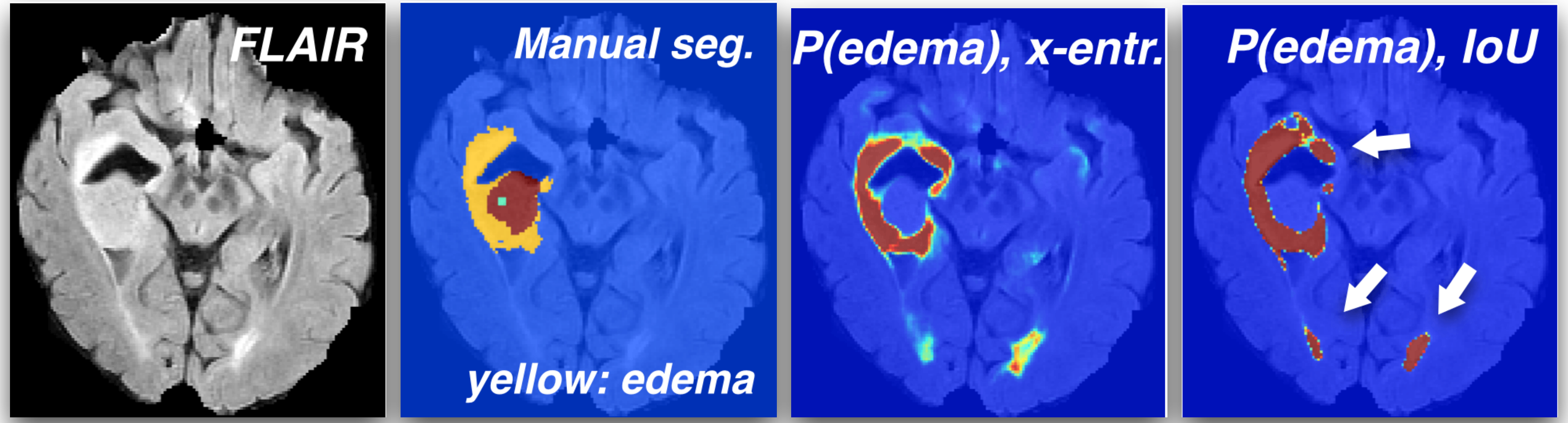
- **The good:**
  - There are some good/promising CNN-based segmentation approaches (Dee
- **The**
  - A lo
  - Arc
  - Arc
- **The ugly:**
  - Chosen model & config may be suboptimal for other data/task
  - Results and conclusions of analysis are strongly biased

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: $P\left(Y|X\right)$

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

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***:

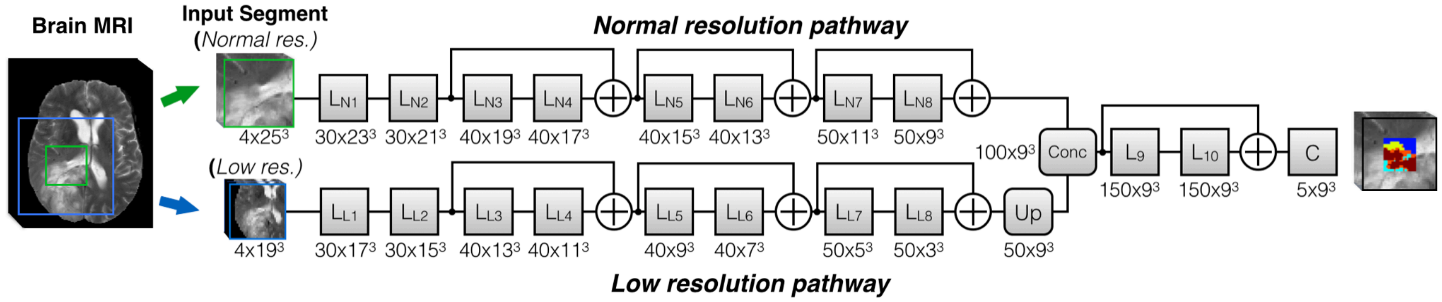
$$P\left(Y|X\right) = \sum_{\forall m \in M} P\left(Y, M=m|X\right) = \sum_{\forall m \in M} P\left(Y|X, M=m\right) P\left(M=m\right)$$

EMMA approximate the joint by ensembling individual models:

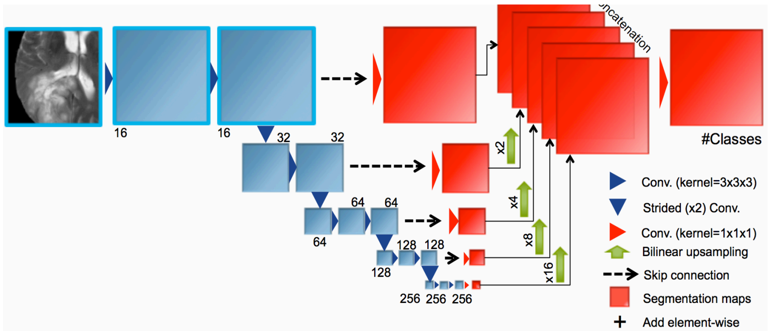
$$\sim P_{EMMA}\left(Y|X\right) = \sum_{\forall m \in M} P\left(Y|X;\theta_m,m\right) \frac{1}{|M|}$$

M: Network architectures

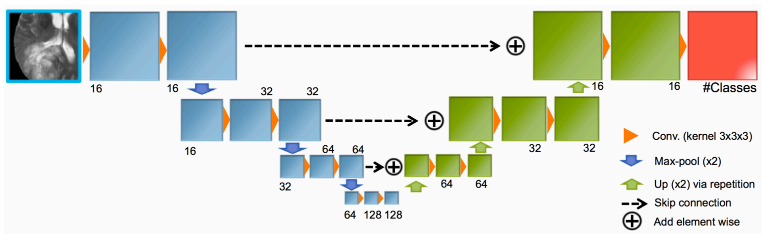
DeepMedic [Kamnitsas 2015, 2016, 2017]



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.

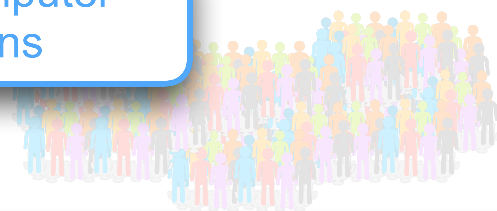
Summary and Conclusions

- Deep learning already plays a crucial role in m
  - Ima
  - Image quantification and analysis

Validation is challenging

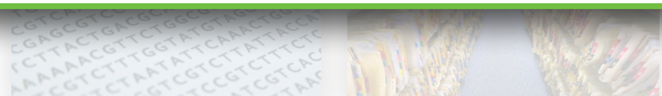
- Applications of deep learning in computer-aided been limited so
  - But there is some (unjustified) hype

Requires collaboration between computer scientists, engineers and clinicians

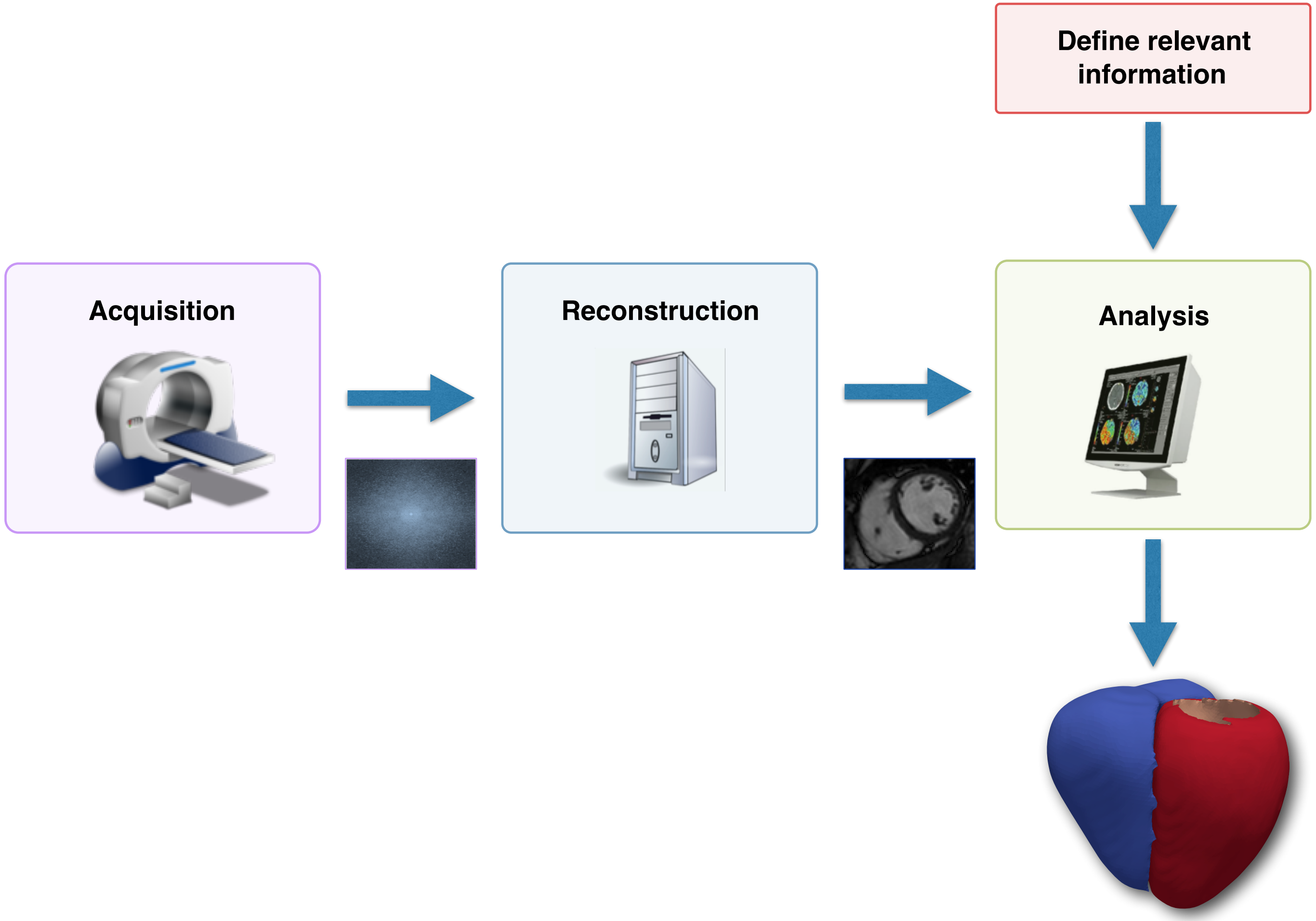


- There is great potential for deep learning for truly intelligent computer-aided diagnosis
  - Learning from unlabelled, large-scale population data
  - Integration of imaging and non-imaging information, e.g. clinical records and genetics

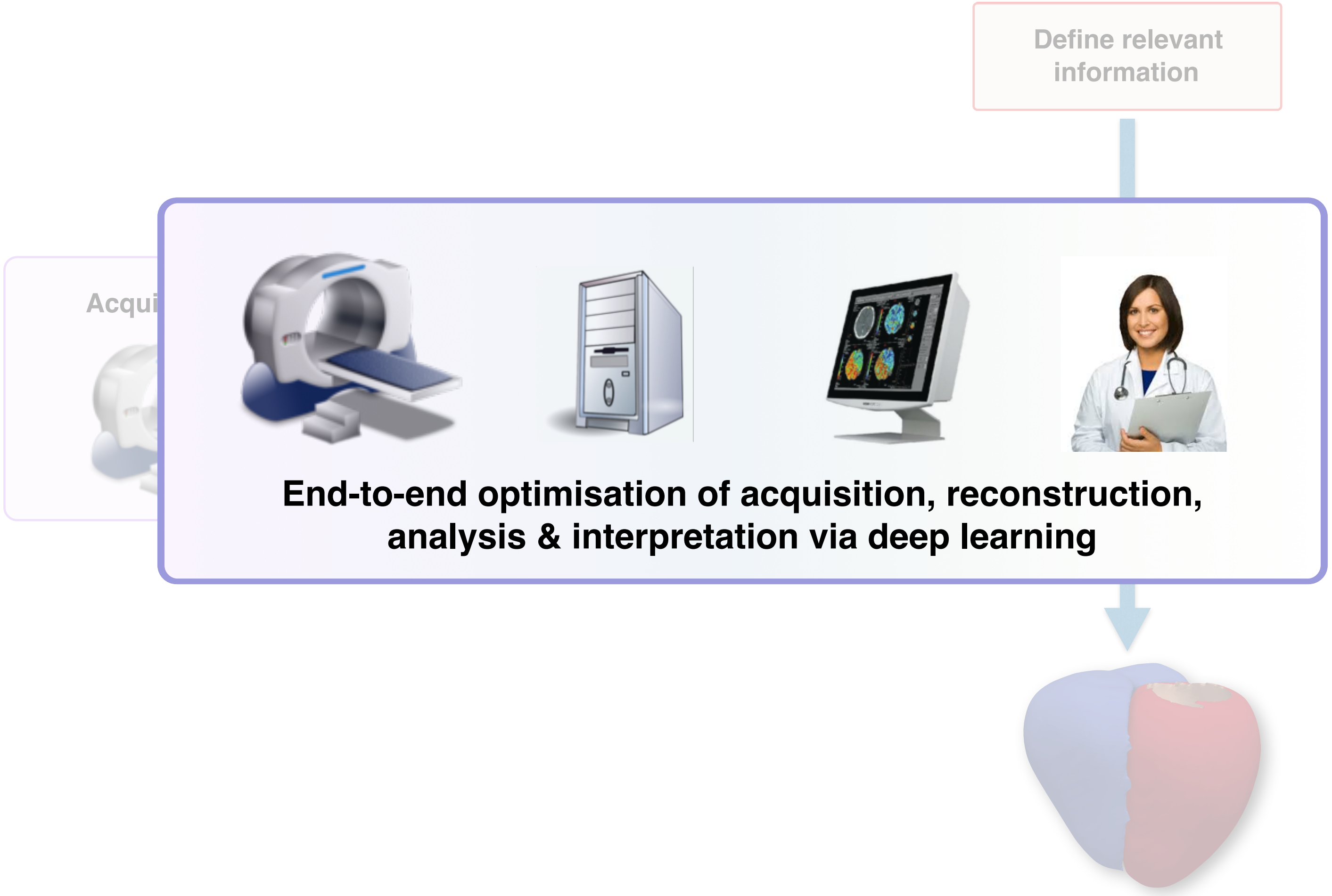
Optimisation of imaging pipeline with respect to clinically useful information



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

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

EPSRC
Engineering and Physical Sciences Research Council

erc
European Research Council

eHealth

SEVENTH FRAMEWORK PROGRAMME

European Commission
Information Society and Media

SmartHeart

PredictND

wellcome trust

NHS
National Institute for Health Research

NVIDIA

Imperial College London