Imperial College London

Learning Anatomical Image Representations for Cardiac Imaging

Ozan Oktay, PhD

Research Scientist @HeartFlow Inc. (CA/USA)
Research Fellow @Imperial College London

Deep Learning for Medical Imaging School
Lyon, April 16th, 2019

Research Background

- Siemens Corporate Research (NJ, USA) (2012-13)
  - Visiting Researcher (Thesis Project)

- Imperial College London (BioMedIA, London, UK)
  - Research Assistant - PhD student (2013 - 2017)
  - Research Associate (2017 - 2018)
  - Honorary Research Fellow (2018 - Present)

- HeartFlow Inc. (CA, USA)
  - Imaging Research Scientist - Scientific Lead (2018 - Present)

- Specialisation:
  - Design and Development of Innovative Machine Learning Solutions for Improved Healthcare Services

Agenda

- *Machine Learning in Medical Image Analysis*
- *Representation Learning in Medical Image Analysis*
- *Inverse Problems and Image Quality Assessment*
- *Building Clinical Solutions (HeartFlow Inc.)*

Machine Learning in Medical Image Analysis

BioMedIA
Imperial College London

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.

NOVEMBER 15, 2017

Stanford algorithm can diagnose pneumonia better than radiologists

Stanford researchers have developed a deep learning algorithm that evaluates chest X-rays for signs of disease. In just over a month of development, their algorithm outperformed expert radiologists at diagnosing pneumonia.

Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening

Nan Wu[b], Jason Phang[b], Jungkyu Park[b], Yiqiu Shen[b], Zhe Huang[b], Masha Zorin[h*], Stanisław Jastrzębski[i], Thibault Févry[b], Joe Katsnelson[f], Eric Kim[a], Stacey Wolfson[a], Ujas Parikh[a], Sushma Gaddam[a], Leng Leng Young Lin[a], Kara Ho[j*], Joshua D. Weinstein[a], Beatriu Reig[a,d], Yiming Gao[a,d], Hildegard Toth[a,d], Kristine Pysarenko[a,d], Alana Lewin[a,d], Jiyon Lee[a,d], Krystal Airola[a], Eralda Mema[a], Stephanie Chung[a], Esther Hwang[a], Naziya Samreen[a], S. Gene Kim[a,d,e], Laura Heacock[a,d], Linda Moy[a,d,e], Kyunghyun Cho[b,c,g], and Krzysztof J. Geras[a,b,e,1]

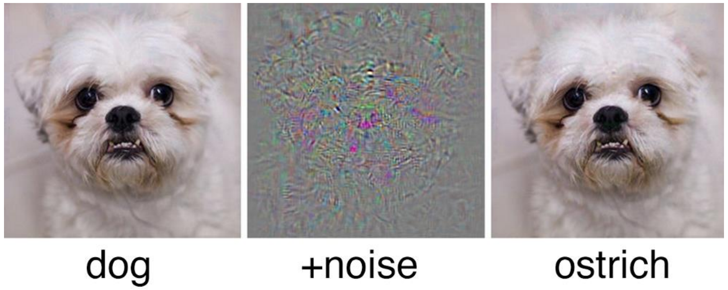
Machine Learning and AI: Are they the same?

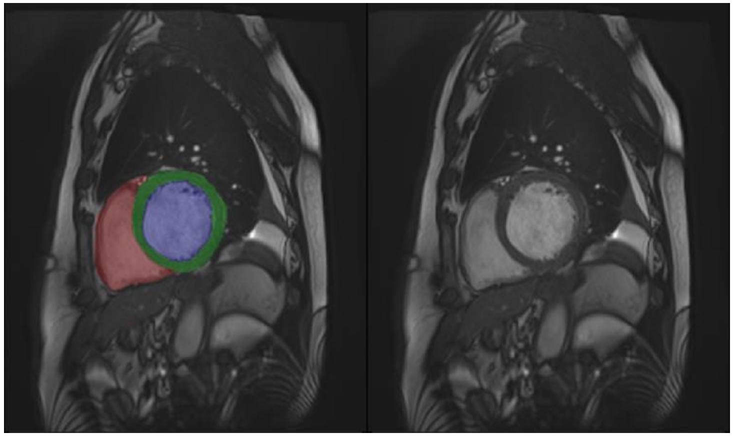
1. What is the difference between AI and Machine Learning?

2. Do neural networks generalise well to unseen data points (e.g. scanner types)? What do these models actually learn?

3. Why do neural networks require large amount of data?

4. Can we fully rely on these algorithms in clinical practice?





Where do we stand?

Machine Learning (ML) models are good at:

- Automating well-defined and constrained (low variation) tasks.
- Reproducible results and fast predictions -> Scale up to thousands of images
- Can display performance close to the average annotator for some tasks -> Label quality

They are a good candidate for assistive clinical workflows as long as they are monitored.

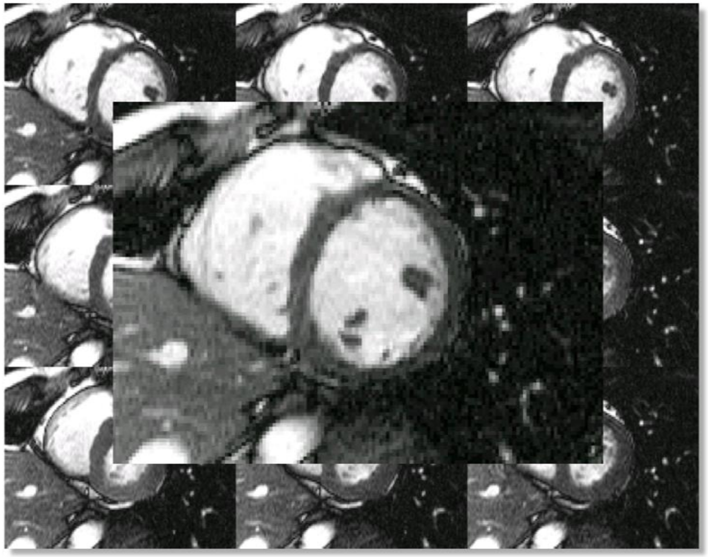
Segmentation of Cardiac Images

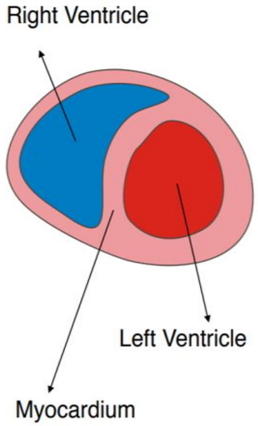
Challenges

- Manual delineation of large number of images can be time-consuming.
- Reproducibility and large inter-observer variability are common issues.

Objectives

- Machine learning algorithm can automate and produce average annotator performance on some tasks.
- It can scale up to thousands of images and assist clinical workflow.





CNN Segmentation Model

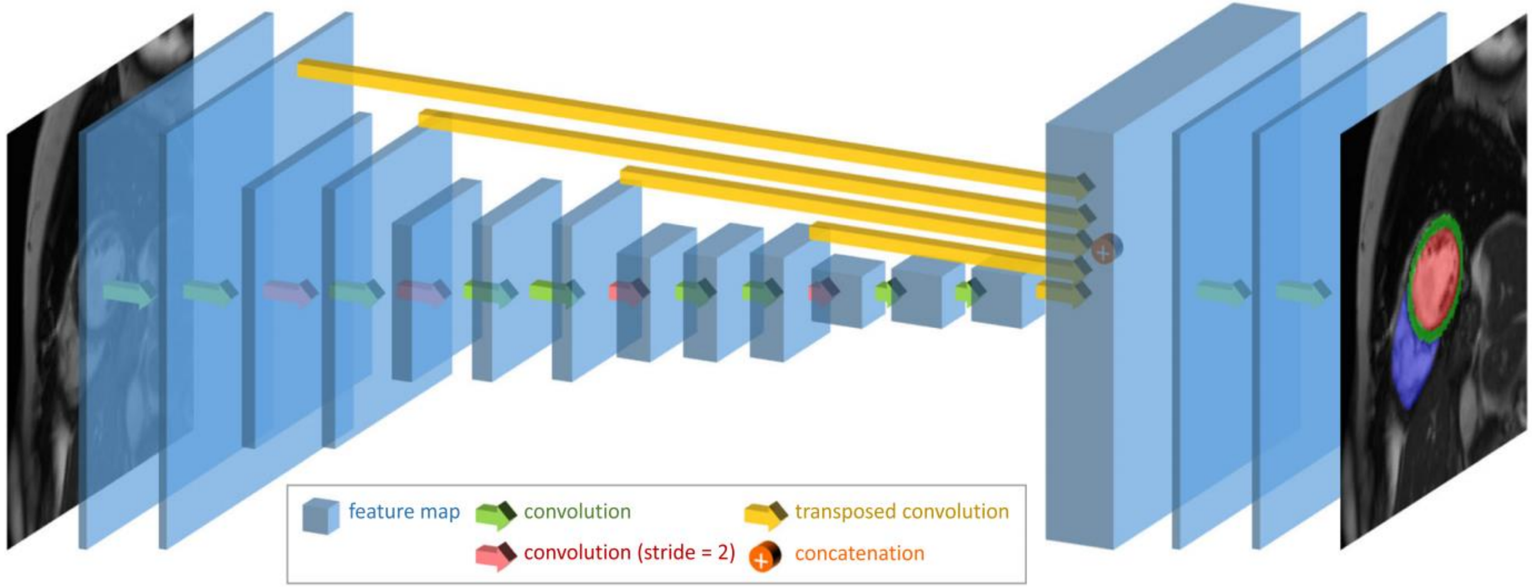


Fig. 1 The network architecture. A fully convolutional network (FCN) is used, which takes the cardiovascular magnetic resonance (CMR) image as input, learns image features from fine to coarse scales through a series of convolutions, concatenates multi-scale features and finally predicts a pixelwise image segmentation

Training dataset: 5000 Cases from the UKBB Dataset.

8 different annotators extracted manual segmentations for these images

[W. Bai et al. JCMR 2018]

Auto-Generated Image Segmentations

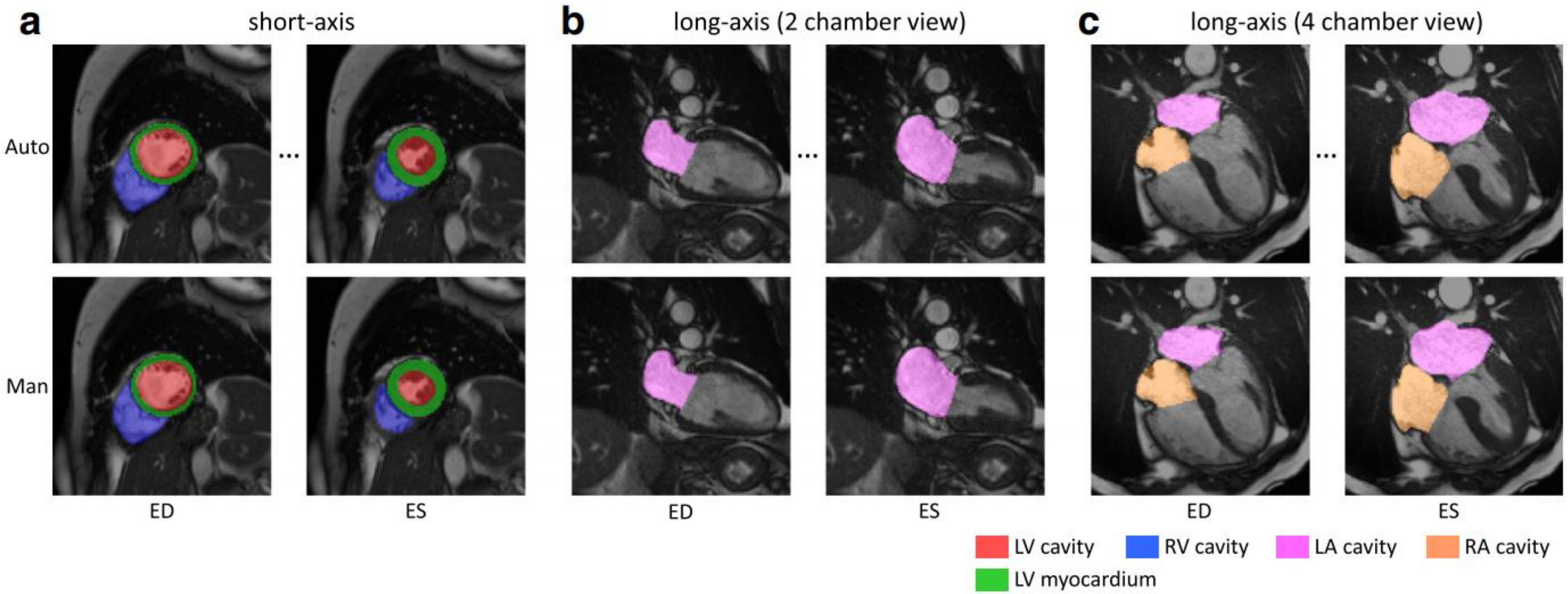


Fig. 3 Illustration of the segmentation results for short-axis and long-axis images. The top row shows the automated segmentation, whereas the bottom row shows the manual segmentation. The automated method segments all the time frames. However, only end-diastolic (ED) and end-systolic (ES) frames are shown, as manual analysis only annotates ED and ES frames. The cardiac chambers are represented by different colours. **a** short-axis. **b** long-axis (2 chamber view). **c** long-axis (4 chamber view)

Variability Between the Annotators and Neural Network Model

BioMedIA Imperial College London

(a) Dice metric				
	Auto vs Manual	O1 vs O2	O2 vs O3	O3 vs O1
	($n = 600$)	($n = 50$)	($n = 50$)	($n = 50$)
LV cavity	0.94 (0.04)	0.94 (0.04)	0.92 (0.04)	0.93 (0.04)
LV myocardium	0.88 (0.03)	0.88 (0.02)	0.87 (0.03)	0.88 (0.02)
RV cavity	0.90 (0.05)	0.87 (0.06)	0.88 (0.05)	0.89 (0.05)
(b) Mean contour distance (mm)				
LV cavity	1.04 (0.35)	1.00 (0.25)	1.30 (0.37)	1.21 (0.48)
LV myocardium	1.14 (0.40)	1.16 (0.34)	1.19 (0.25)	1.21 (0.36)
RV cavity	1.78 (0.70)	2.00 (0.79)	1.78 (0.45)	1.87 (0.74)

Imperial College London

Multi-Input Cardiac Image Super-Resolution using Convolutional Neural Networks

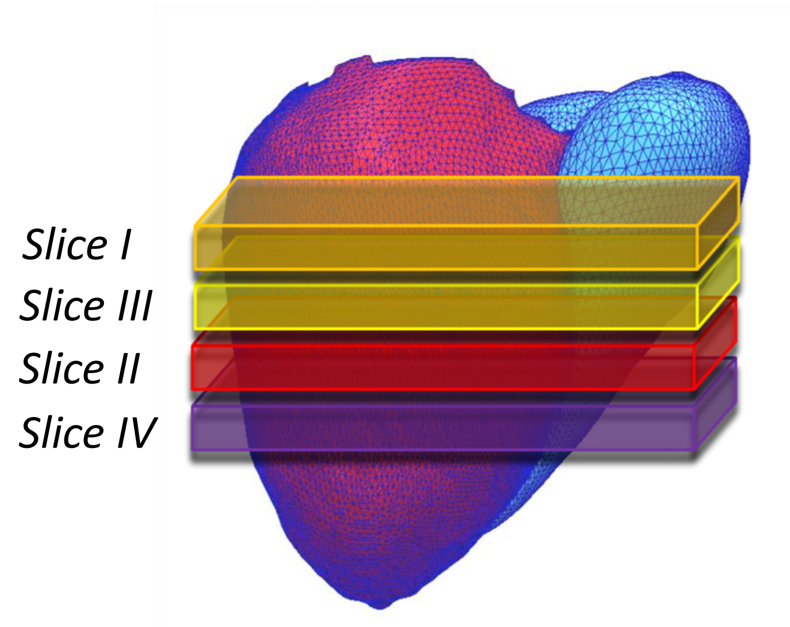
Ozan Oktay, Wenjia Bai, Matthew Lee, Ricardo Guerrero, Konstantinos Kamnitsas, Jose Caballero, Antonio de Marvao, Stuart Cook, Declan O'Regan, and Daniel Rueckert

MICCAI'16 Conference, October 2016, Athens

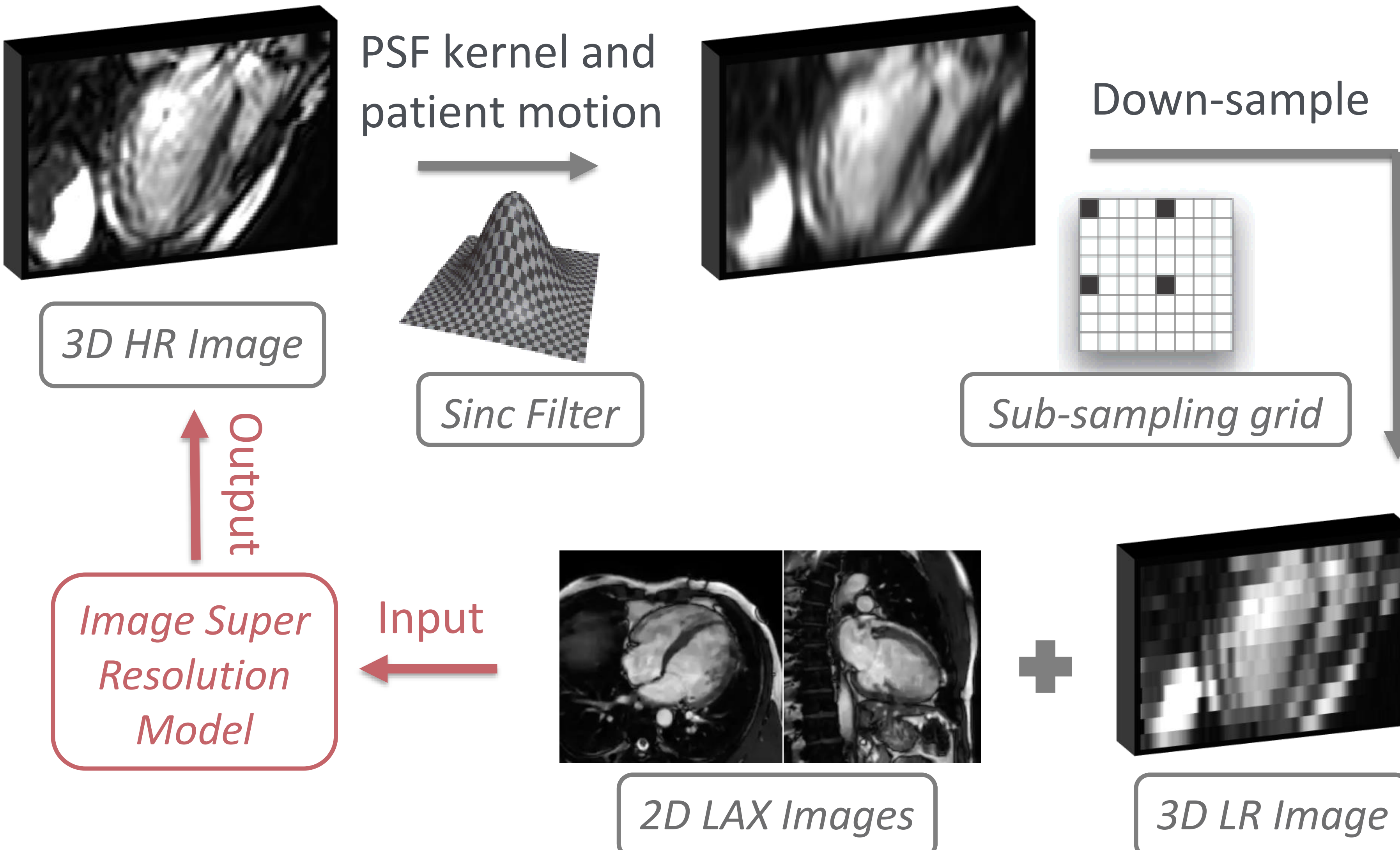
Clinical Motivation

SAX Cardiac MR Image Acquisition

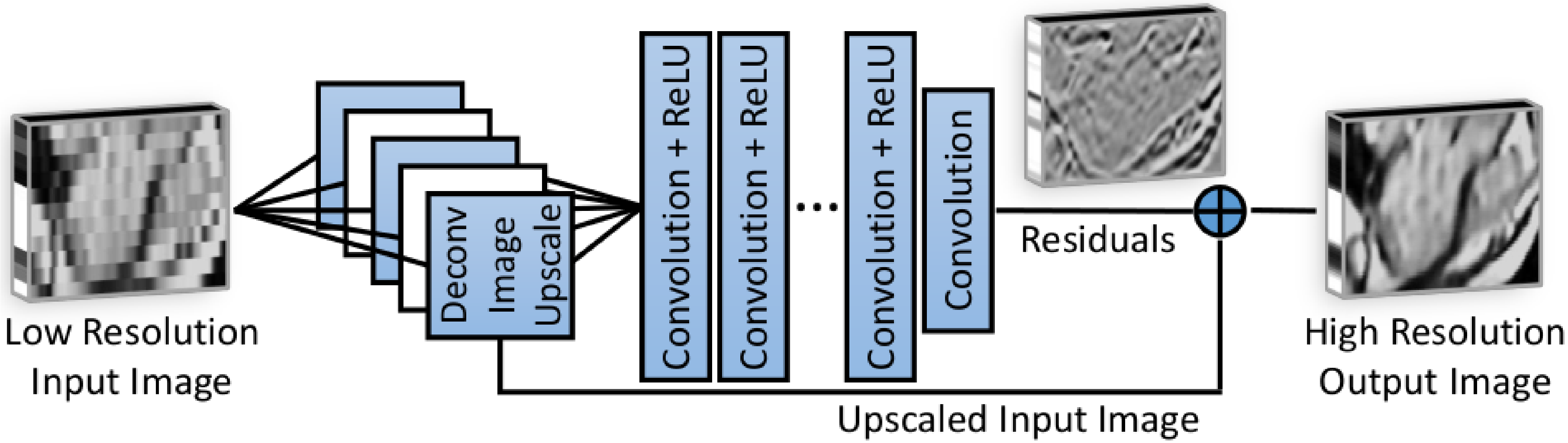
- Large slice thickness (8-10 mm)
    - Due to constrains on SNR, acquisition and breath-hold time
- It hampers subsequent image analysis and quantitative measurements.



Low and High Resolution Images



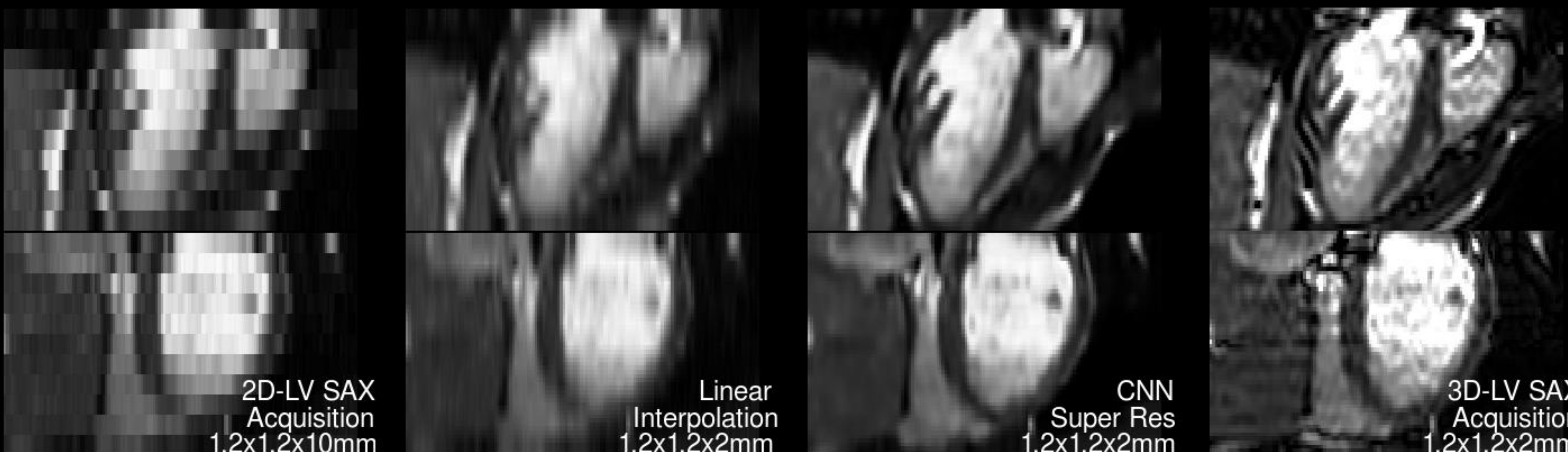
Proposed 3D-SR Model (Single-Image)



Components of the model

- 3D Convolution and Deconvolution (inverse convolution) Kernels
- Rectified Linear Units (ReLUs)
- Regression Based Cost Function (Smooth L1-Norm)
- Input (2D Stack-LR) and Output (3D-HR) Images

Image Quality Assessment

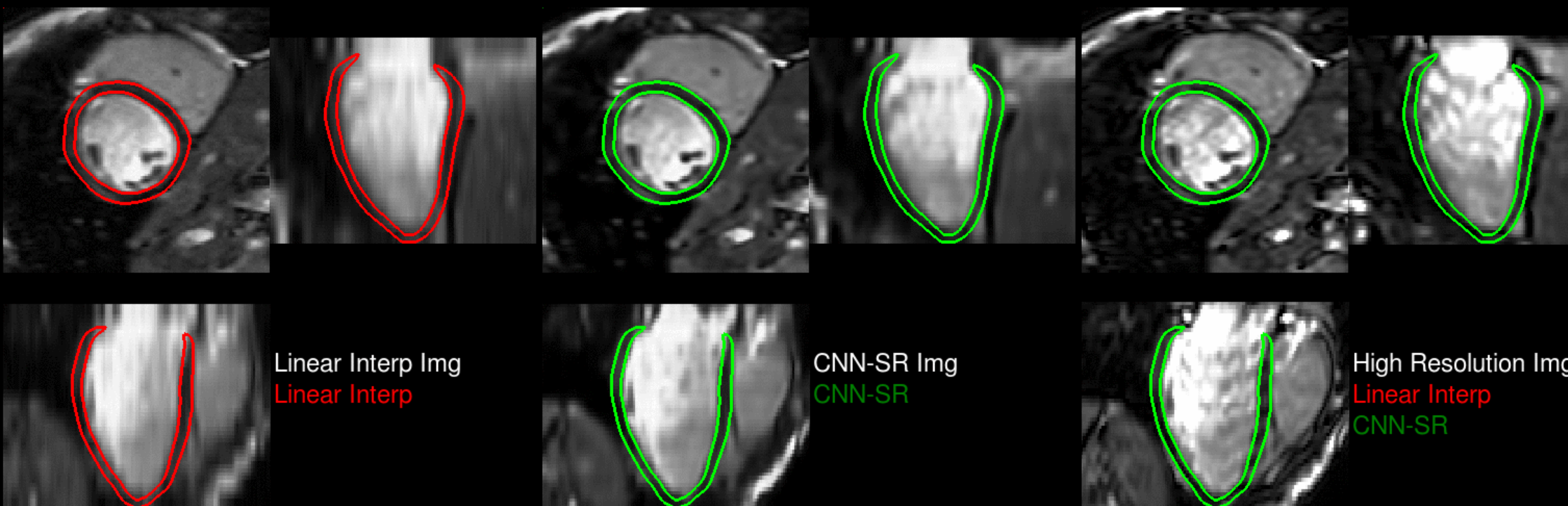


Low Resolution Input Image	Linear Interpolation	The Proposed Method	High Resolution Ground-truth

Upsampling x5

Inference Time: 6-8 Seconds for image size (140x140x10)

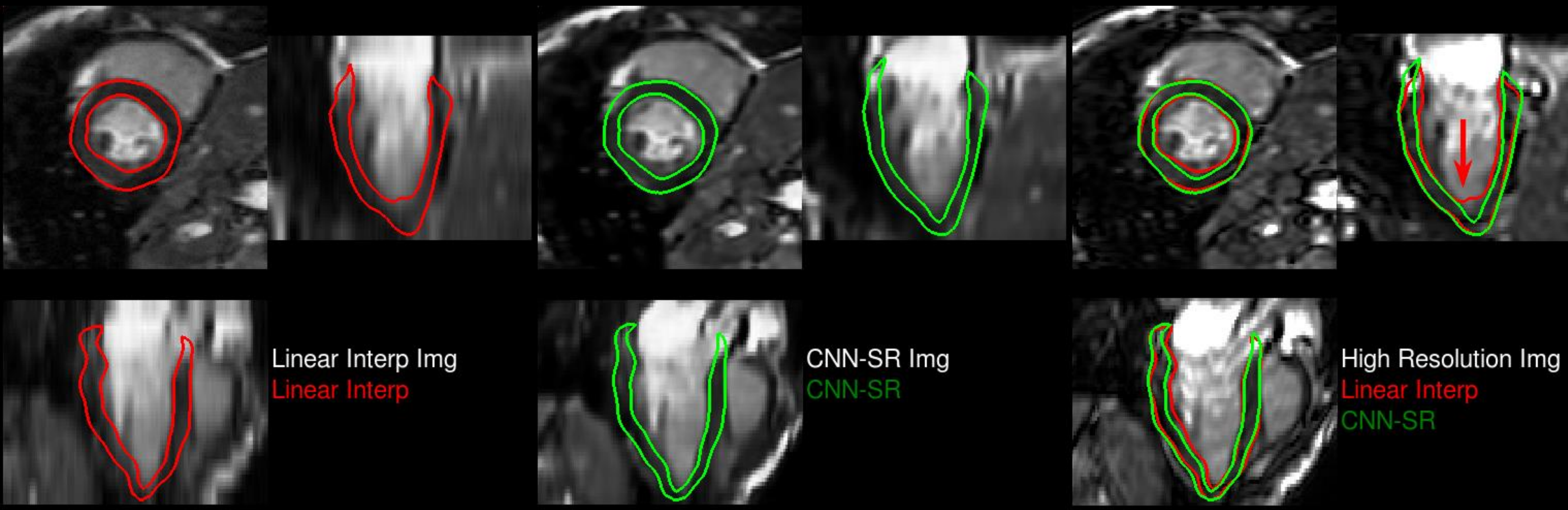
Motion Tracking Experiments (SR is used as a preprocessing method)



Surface to Surface Distance (Linear vs HR) 5.50 mm

Surface to Surface Distance (Proposed vs HR) 4.73 mm

Motion Tracking Experiments (SR is used as a preprocessing method)



Imperial College London

Anatomically Constrained Convolutional Neural Networks (ACNN): Application to Image Enhancement and Segmentation

Ozan Oktay, Enzo Ferrante, Konstantinos Kamnitsas, Wenjia Bai, Jose Caballero, Mattias Heinrich, Stuart Cook, Antonio de Marvao, Declan O'Regan, Bernhard Kainz, Ben Glocker, and Daniel Rueckert

IEEE TMI, August 2017

Research Motivation

Analysis of Neural Networks

I. Model parameterization
II. Model capacity / receptive field
III. Loss function / objective

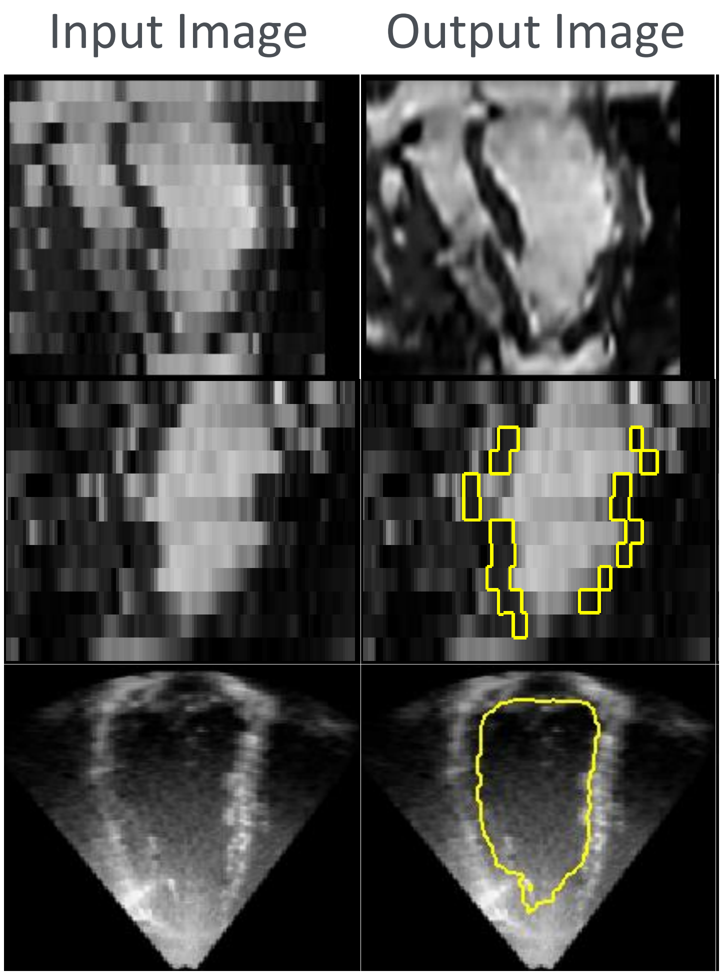
Standard Loss Functions

I. X-Entropy loss function

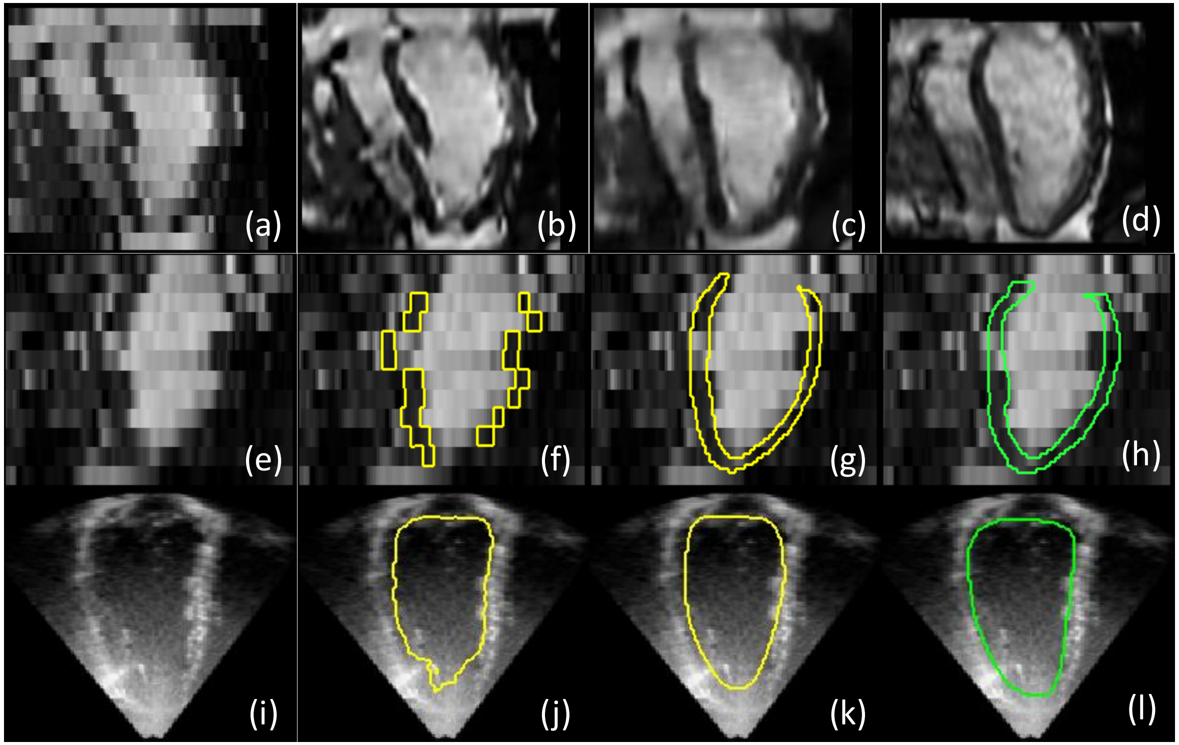
$$L_x = -\sum_{i \in \mathcal{S}} \sum_{c=1}^{C} \log \left(\frac{e^{f(c,i)}}{\sum_j e^{f(j,i)}} \right)$$

II. L2 or Smooth L1 loss function

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



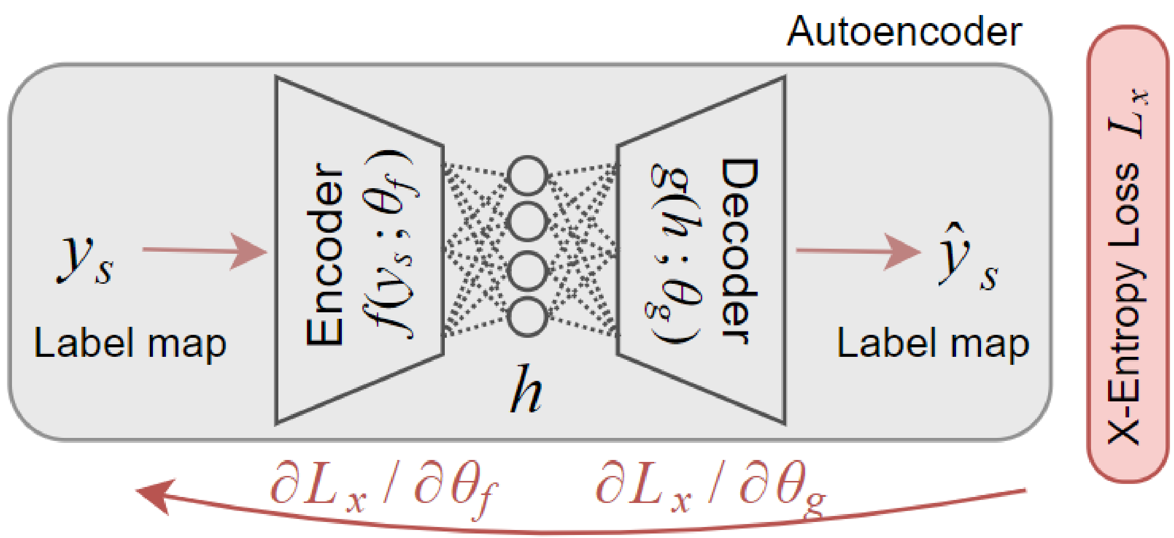
Research Objective



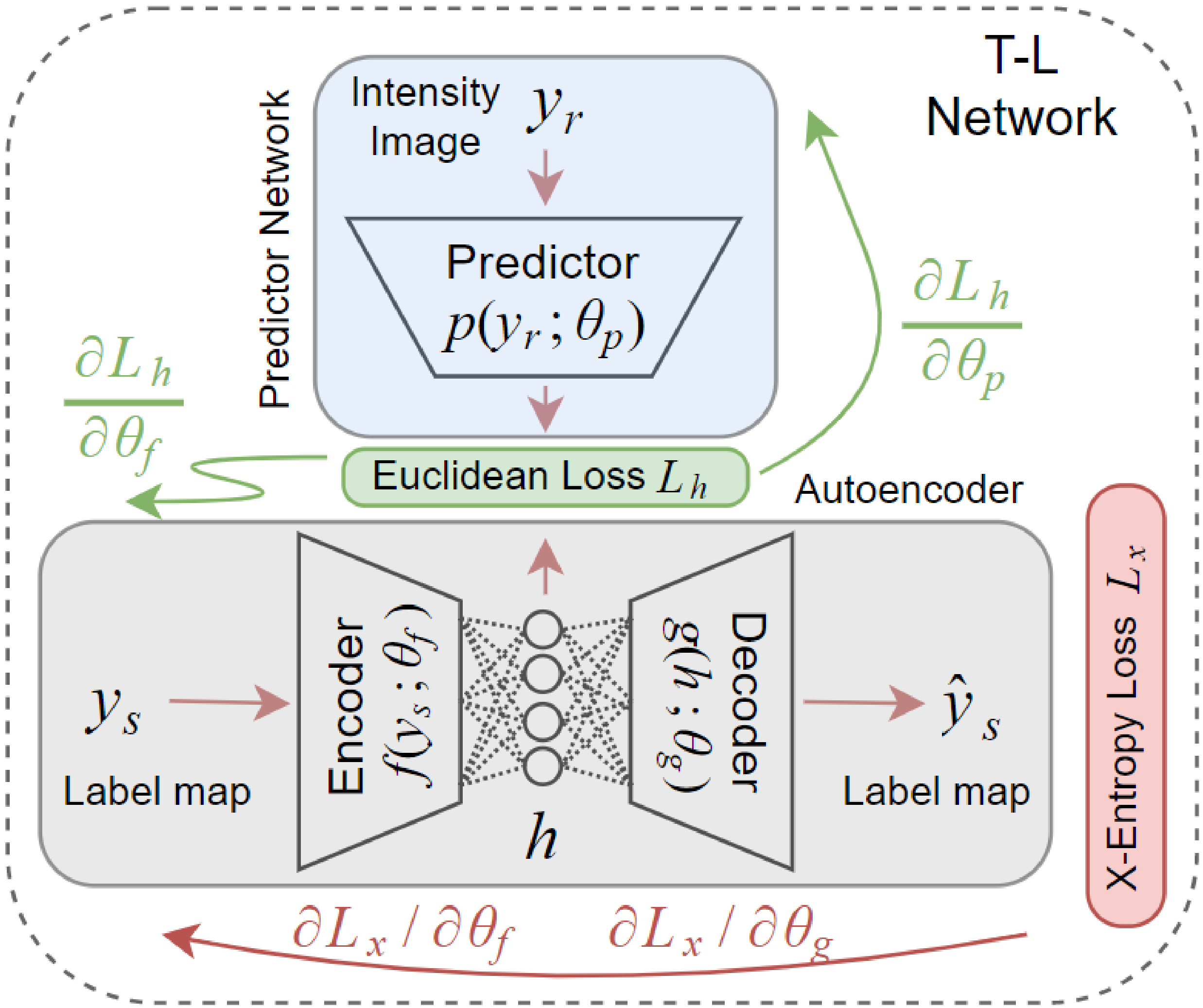
Input Image | Baseline Approaches | Proposed ACNN | Ground-truth

1. Can we teach our models the underlying anatomical priors (eg shape) ?
2. A new global training objective to teach CNN models

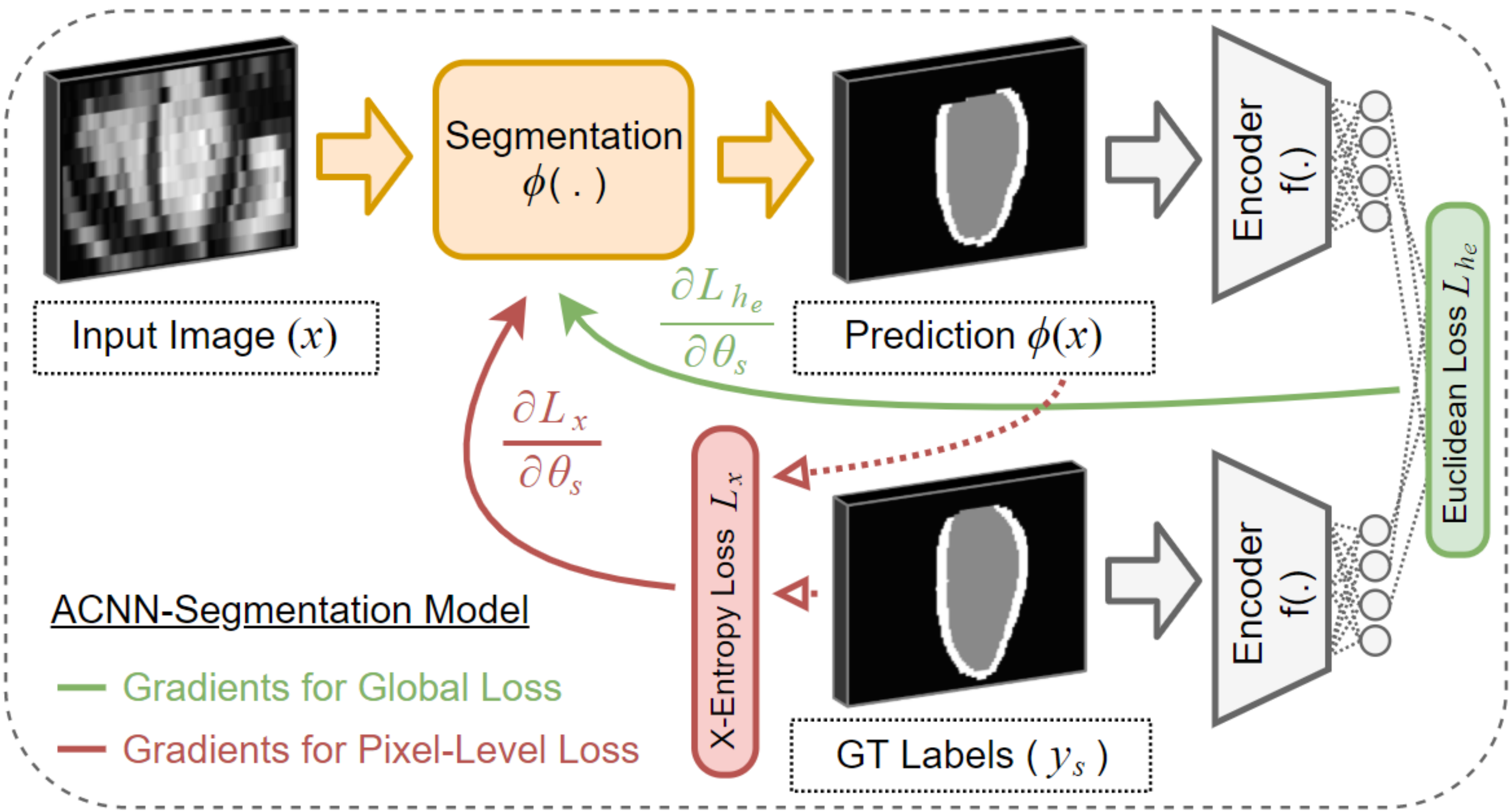
Standard Auto-Encoder Model



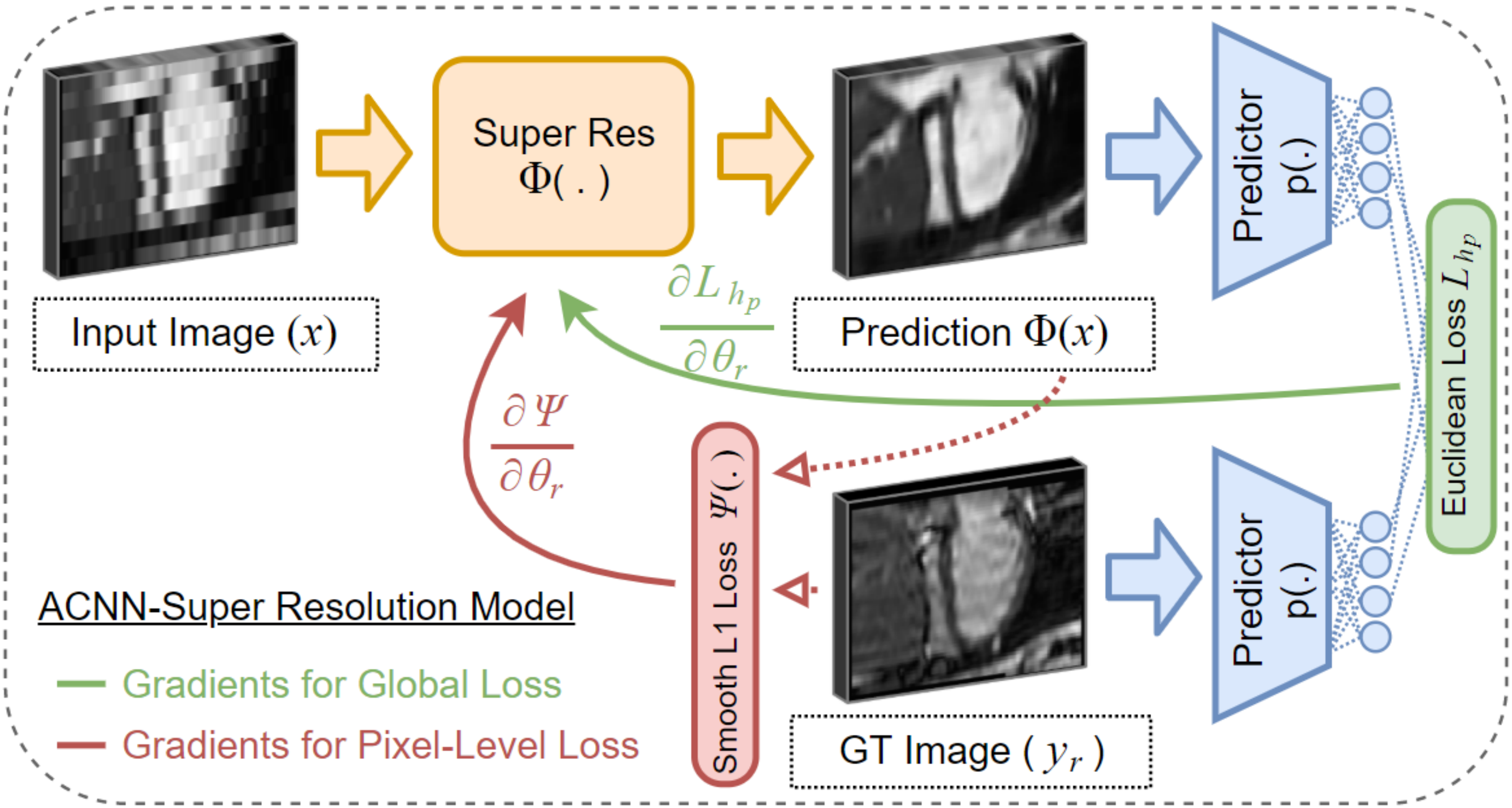
The proposed T-L Network



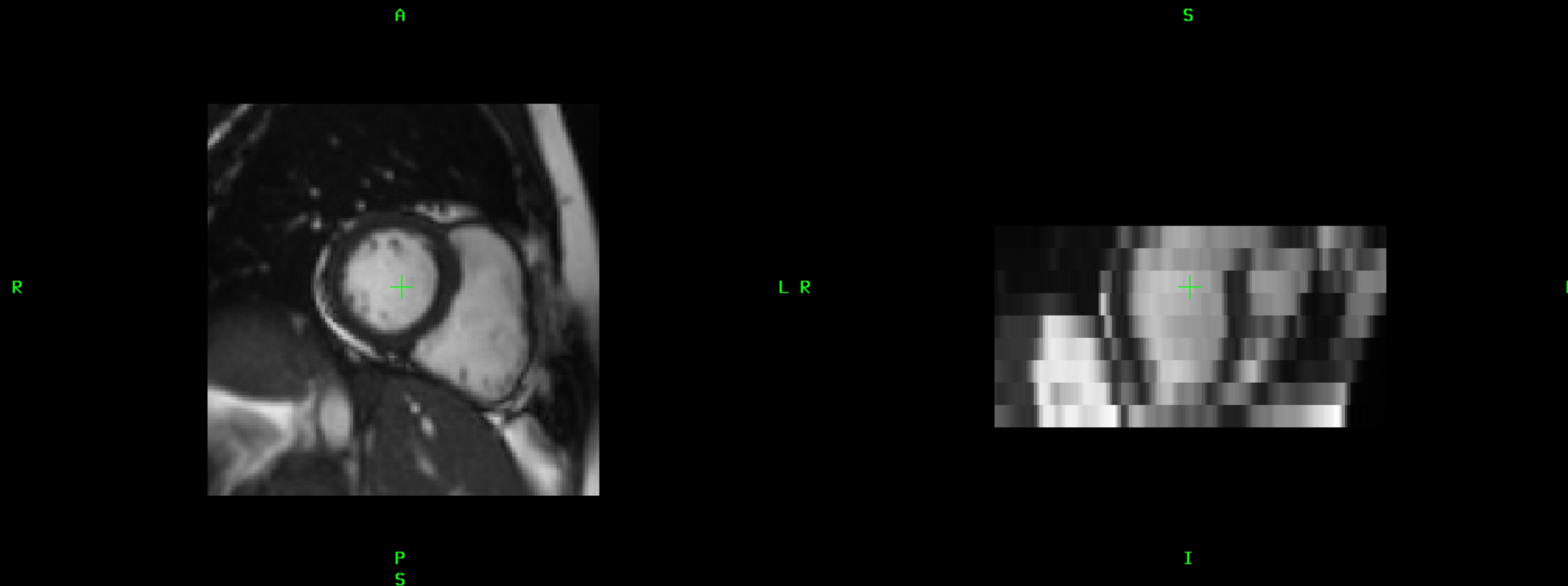
Proposed ACNN - Segmentation Model

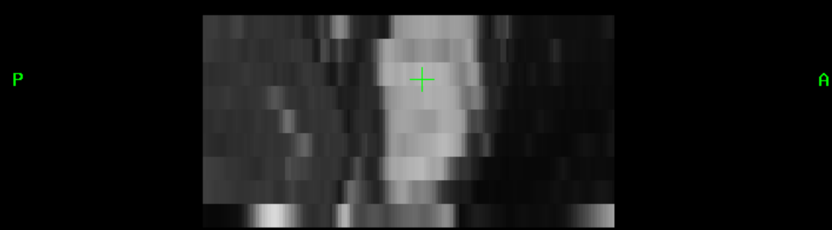


Proposed ACNN – Super Res Model



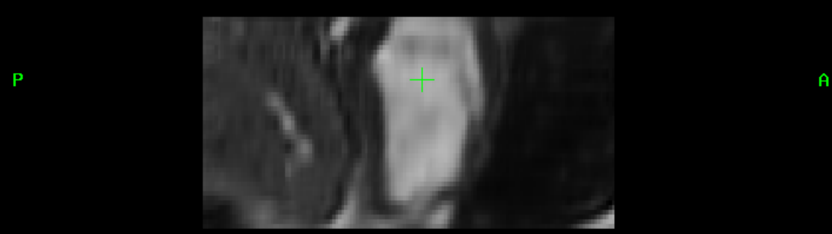
Cardiac MR Super-Resolution Experiments





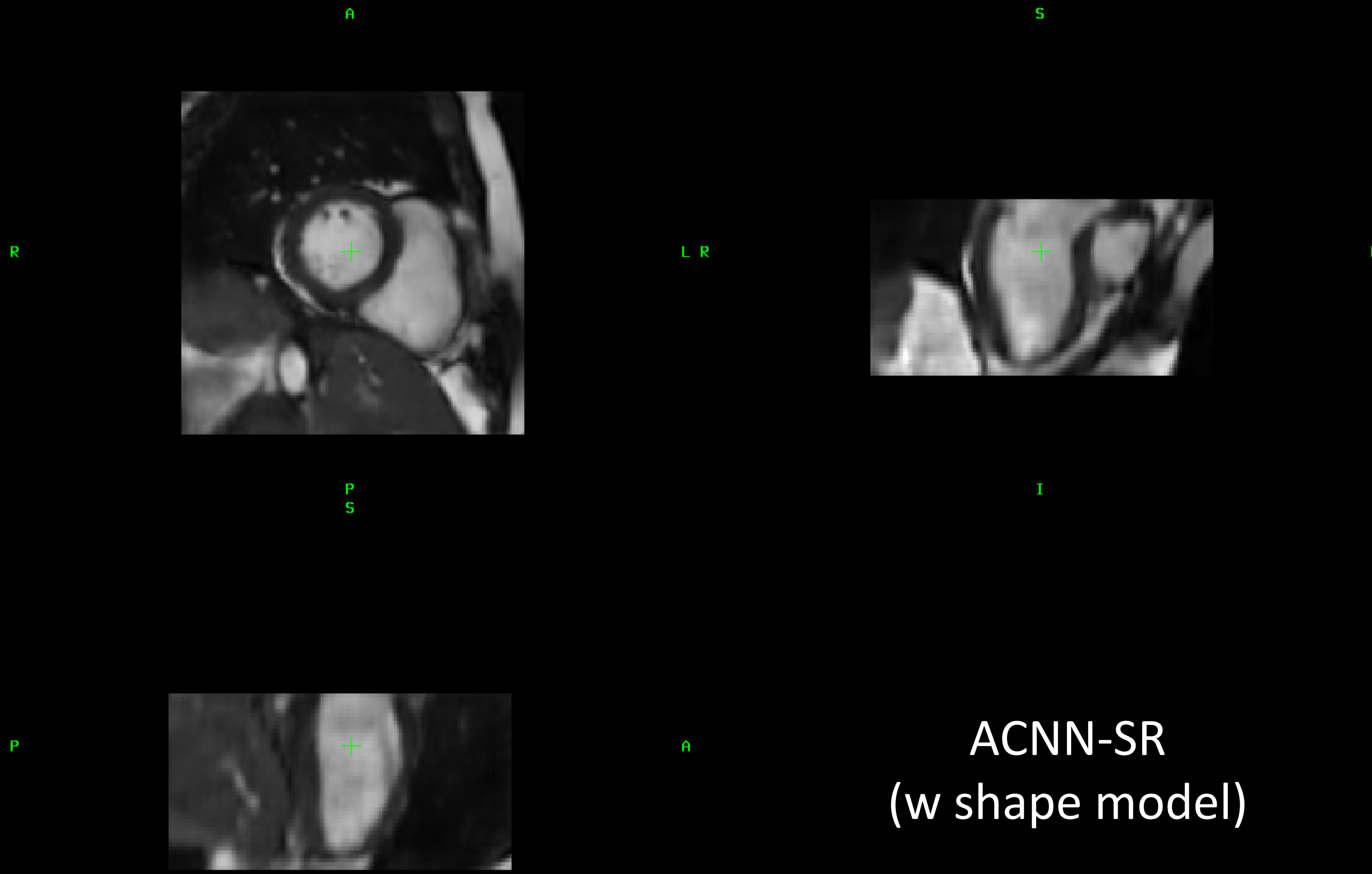
Input Low Resolution Image

Cardiac MR Super-Resolution Experiments



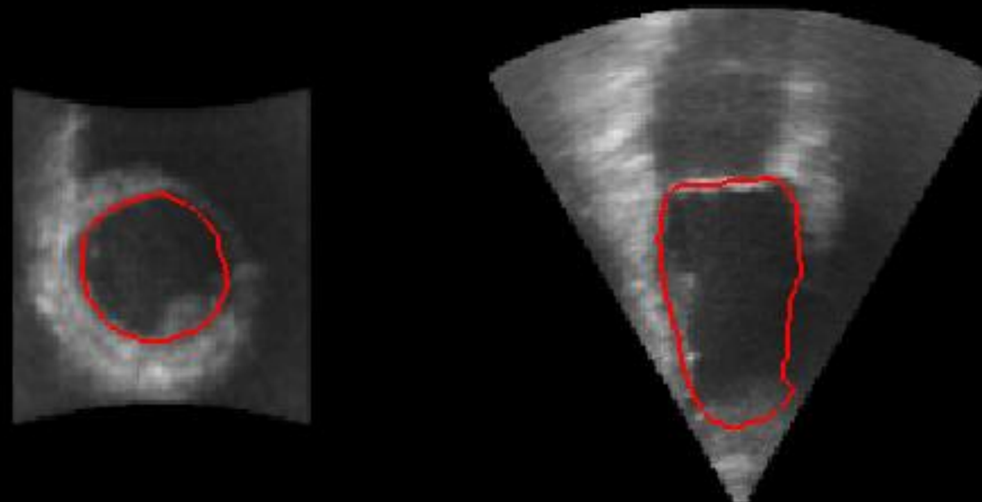
CNN Super-Resolution Trained with Motion-Augmentation

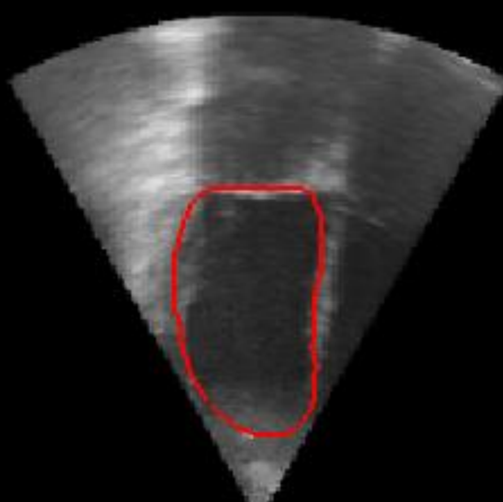
Cardiac MR Super-Resolution Experiments



ACNN-SR
(w shape model)

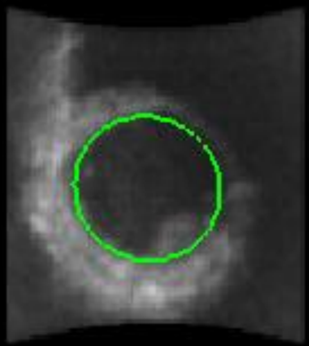
3D-US Segmentations Results

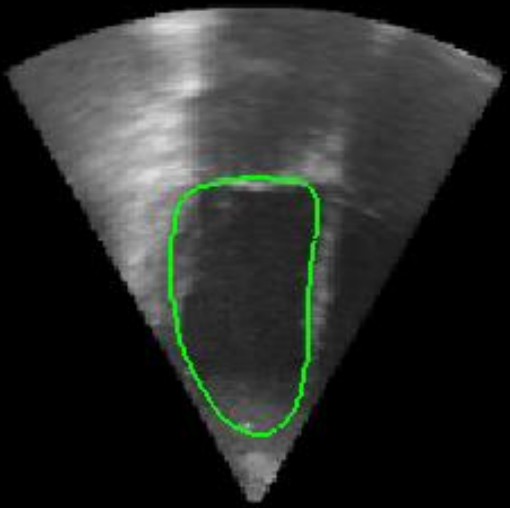




mean dis: 2.94
haus dis: 14.34
dice scr: 0.89

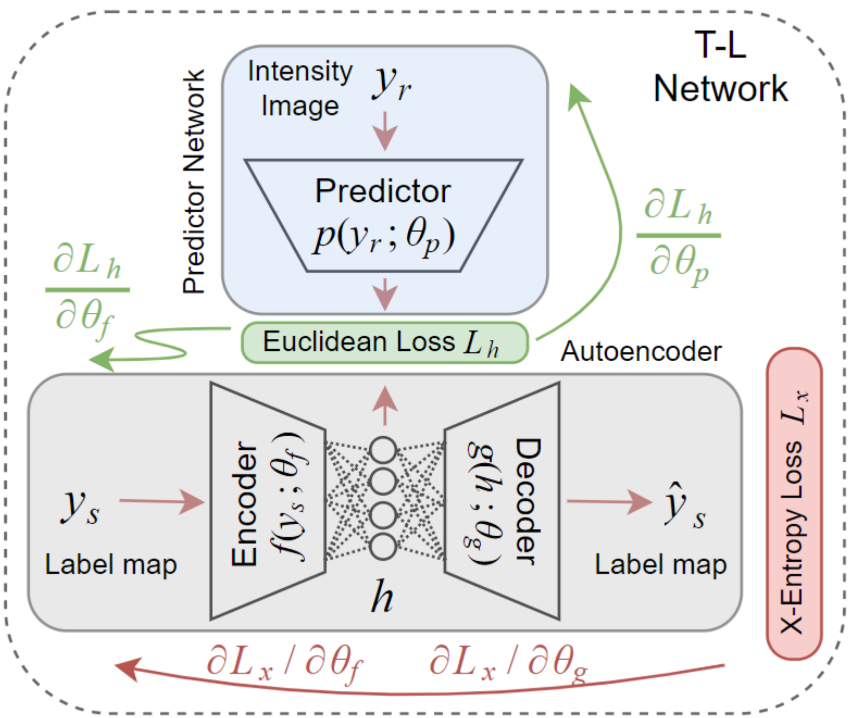
3D-US Segmentations Results



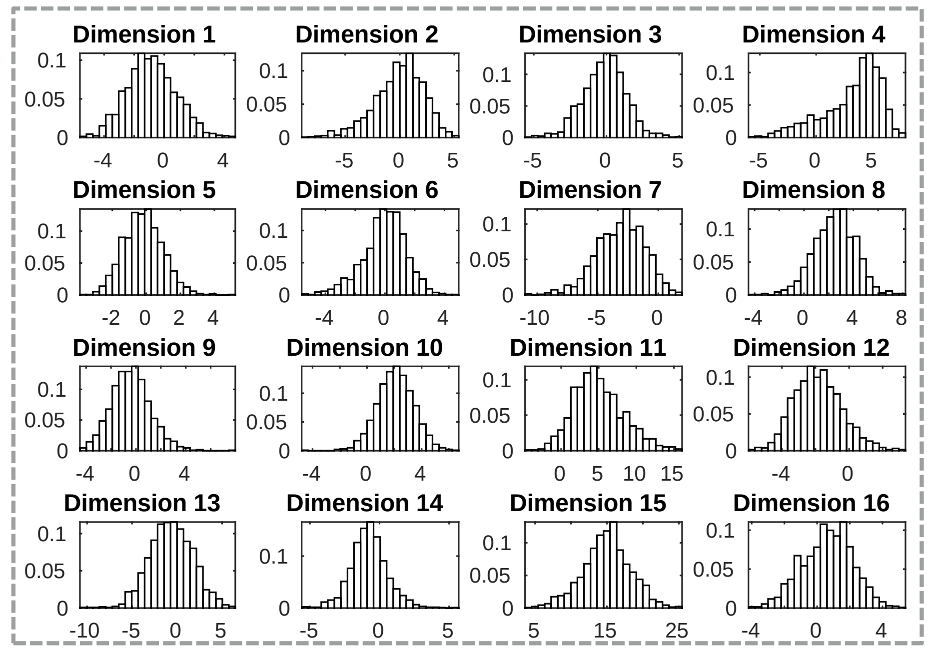


mean dis: 1.74
haus dis: 6.33
dice scr: 0.9

Learned Hidden Representations

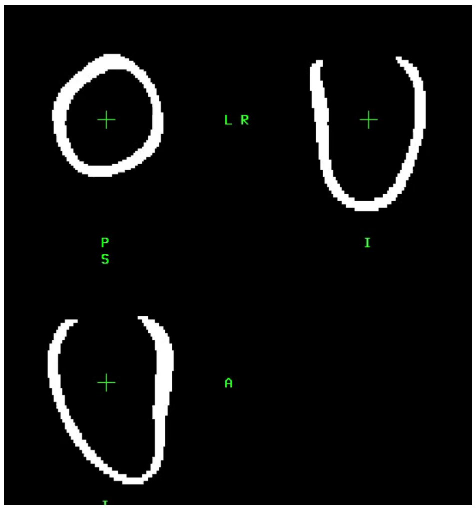


The Proposed Regularization Model

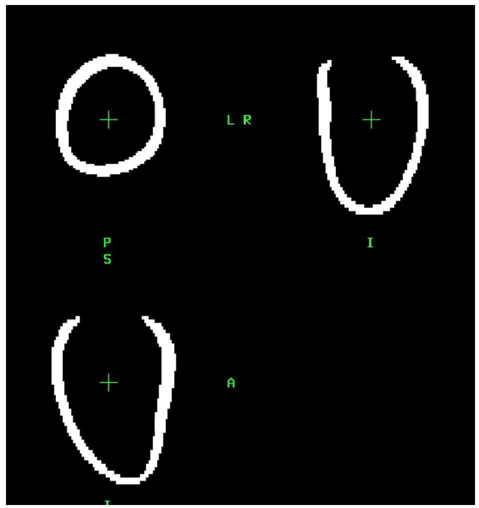


Histogram of the Learned Codes

Learned Hidden Representations



Mean±2Std
(Code #1)



Mean±2Std
(Code #2)

PCA Codes vs T-L Codes

I. Pathology classification
   » Healthy Subjects
   » Dilated Cardiomyopathy
   » Hypertrophic Cardiomyopathy
II. Classification accuracy
   » PCA: 83.3%
   » T-L: 91.6%
   » 60 CMR Sequences

Learned representations can be used to:

I. Predict Clinical Indices (Age, Blood Pressure, Myocardial Mass, etc ..)
II. Genetic Studies / Understanding the cardiac related pathologies

Imperial College London

Learning Based Quality Control for Cardiac MR Images

Giacomo Tarroni, Ozan Oktay, Wenjia Bai, Andreas Schuh, Hideaki Suzuki, Jonathan Passerat-Palmbach, Antonio de Marvao, Declan P. O'Regan, Stuart Cook, Ben Glocker, Paul M. Matthews, Daniel Rueckert

IEEE TMI, November 2018

Automated MR Image Quality Assessment

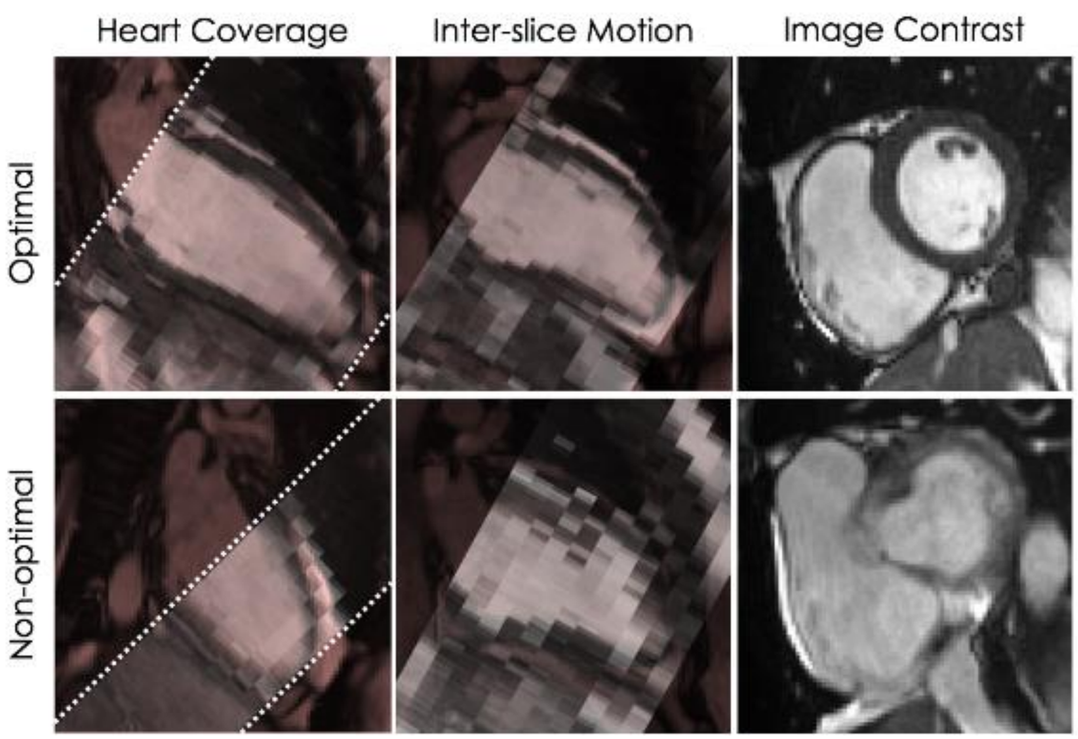
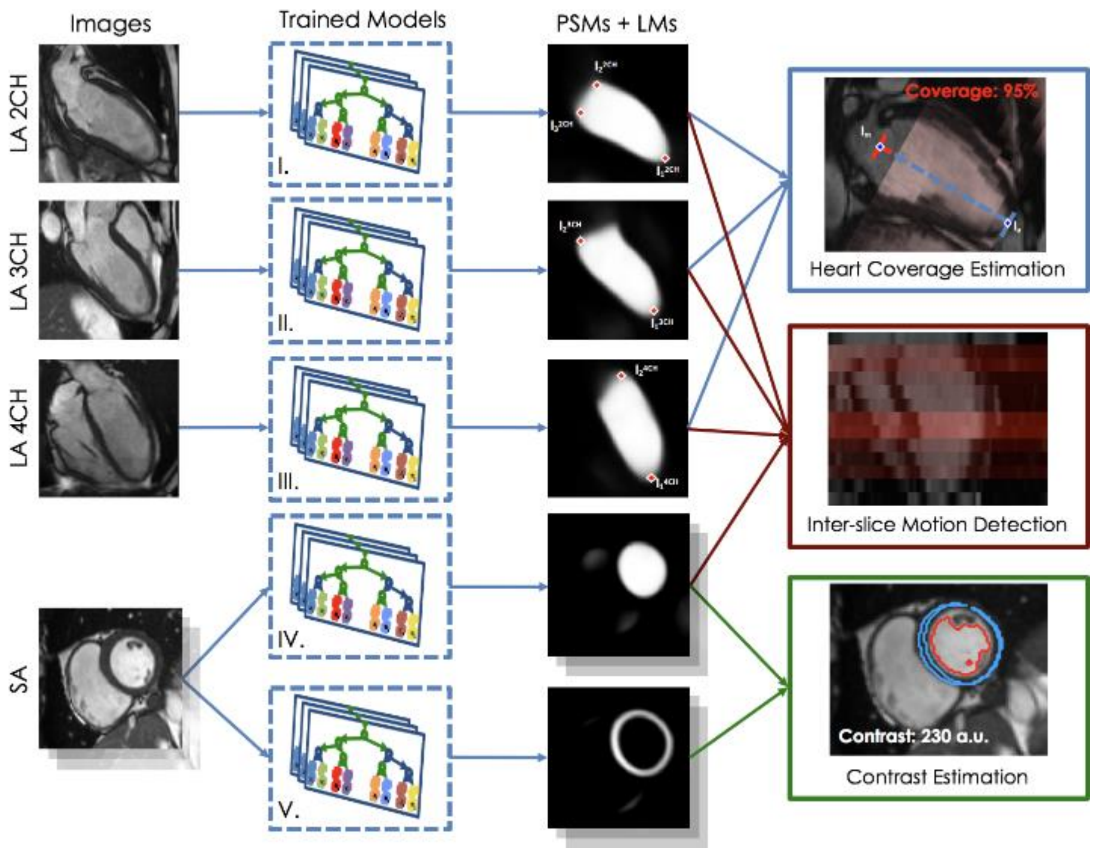


Image Quality Issues

- Affecting short-axis cardiac MR image acquisition.
- Fully-automated quality control pipeline for cardiac MRI, tested on 3000 cases from the UK Biobank study.

Automated MR Image Quality Assessment



Decision tree model automatically quantifies:

- Inter-slice misalignment.
- Heart-coverage rate.
- Image contrast.

[G. Tarroni, O. Oktay et al. IEEE TMI 2018]

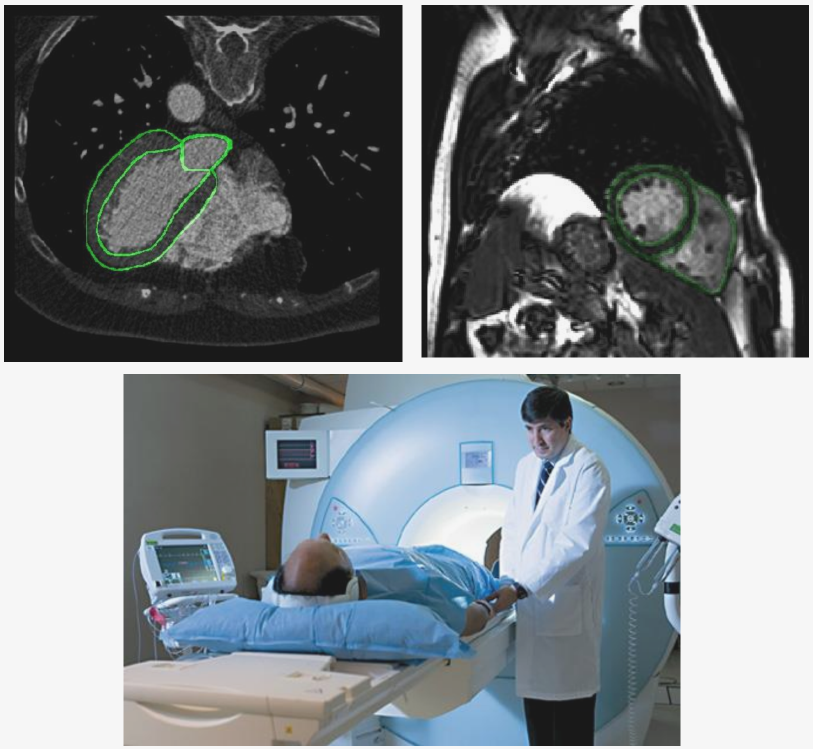
Imperial College London

Structured Decision Forests For Multi-modal Ultrasound Image Registration

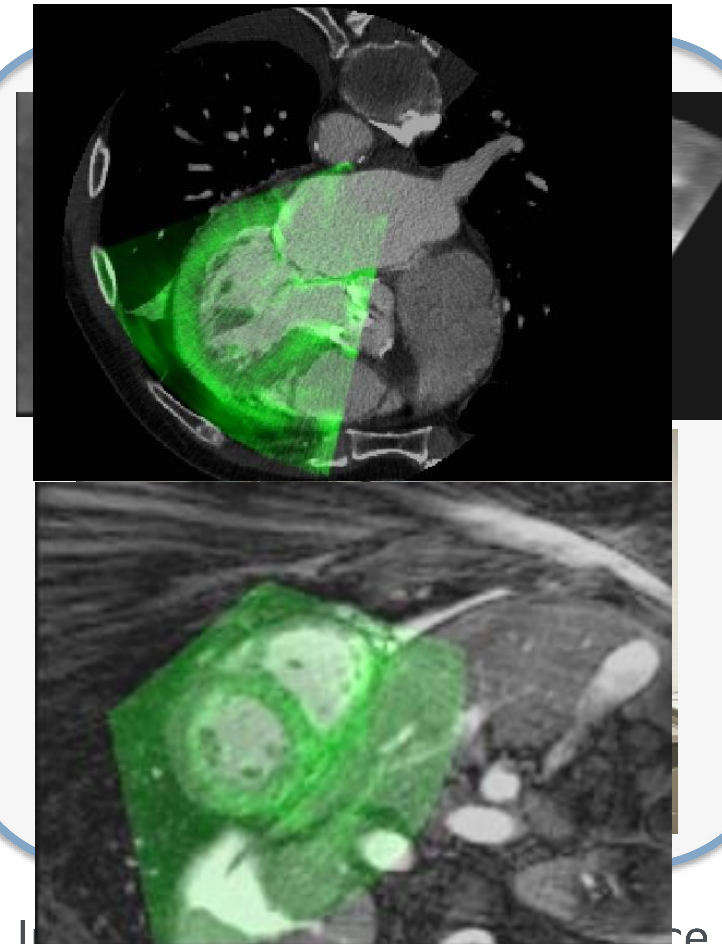
Ozan Oktay, Andreas Schuh, Martin Rajchl, Kevin Keraudren, Alberto Gomez, Mattias Heinrich, Graeme Penney, and Daniel Rueckert

MICCAI'15 Conference, October 2015, Munich

Image Guided Cardiac Interventions



Pre-Operative Stage CT and
MR Image Acquisitions



Spatially Aligned Pre and
Intra Operative Images

Advantages of Probabilistic Edge Maps

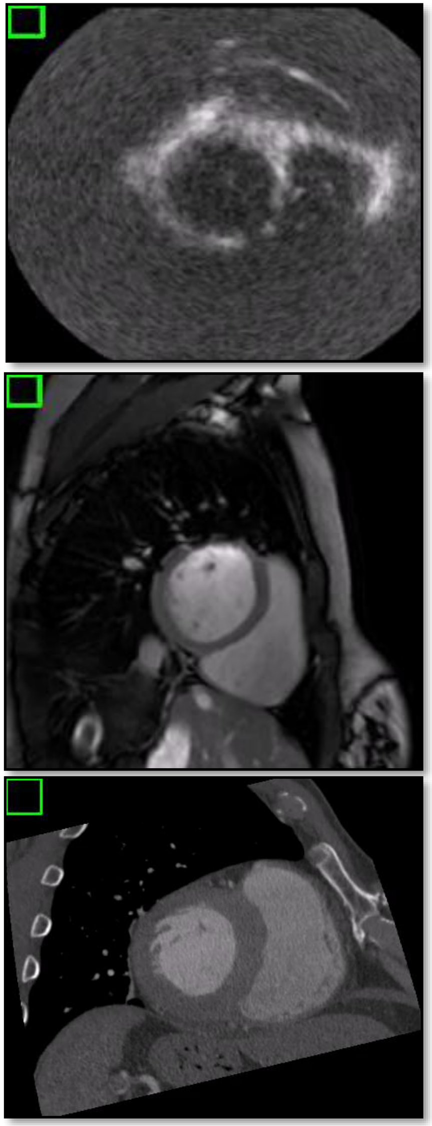
I. Modality independent (e.g. CT, MRI, US)

II. Computationally efficient (20s per image)

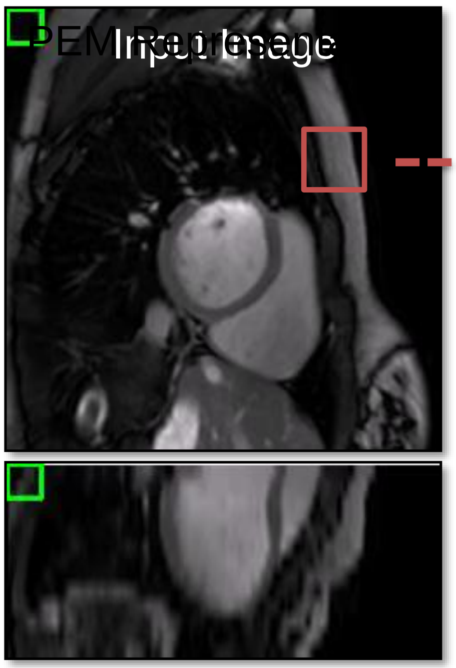
III. Target organ specific image registration

IV. Accurate and smooth anatomical representation

V. Same training and testing configuration is applied to all three modalities.



Structured Decision Forest



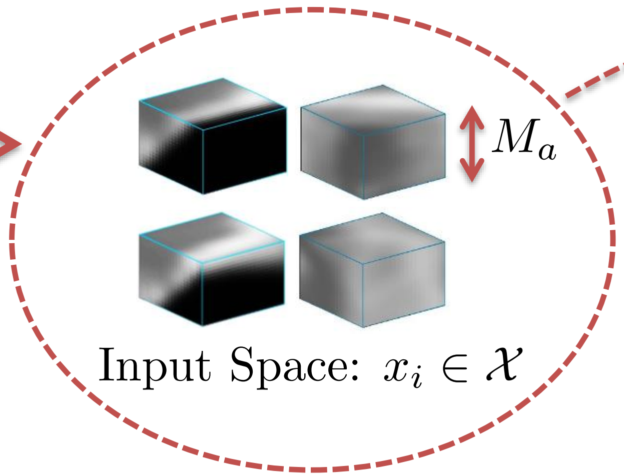
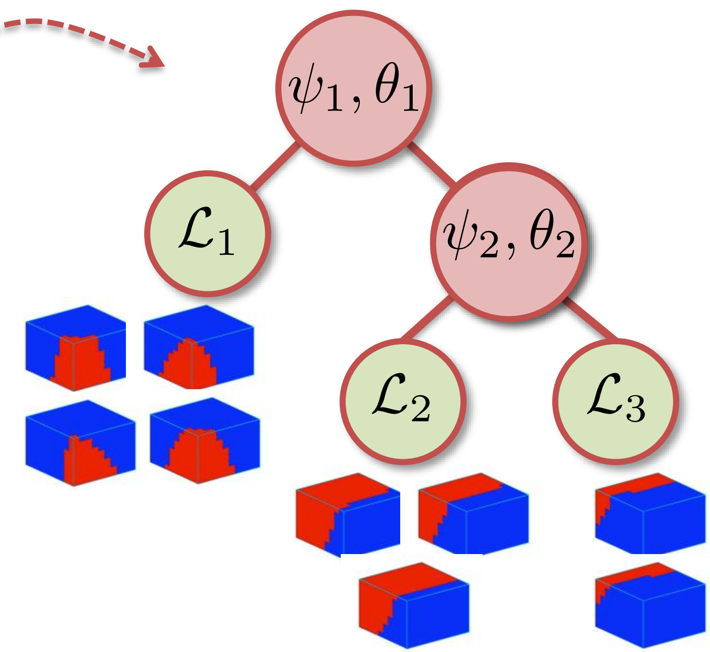
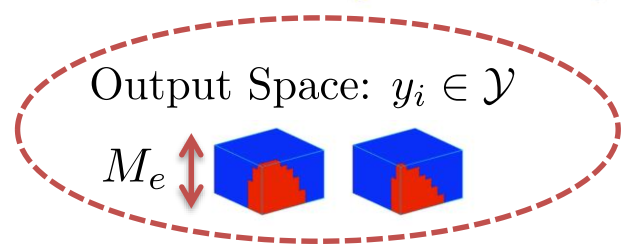


Image features

Structured Decision Tree





Output edge patch labels

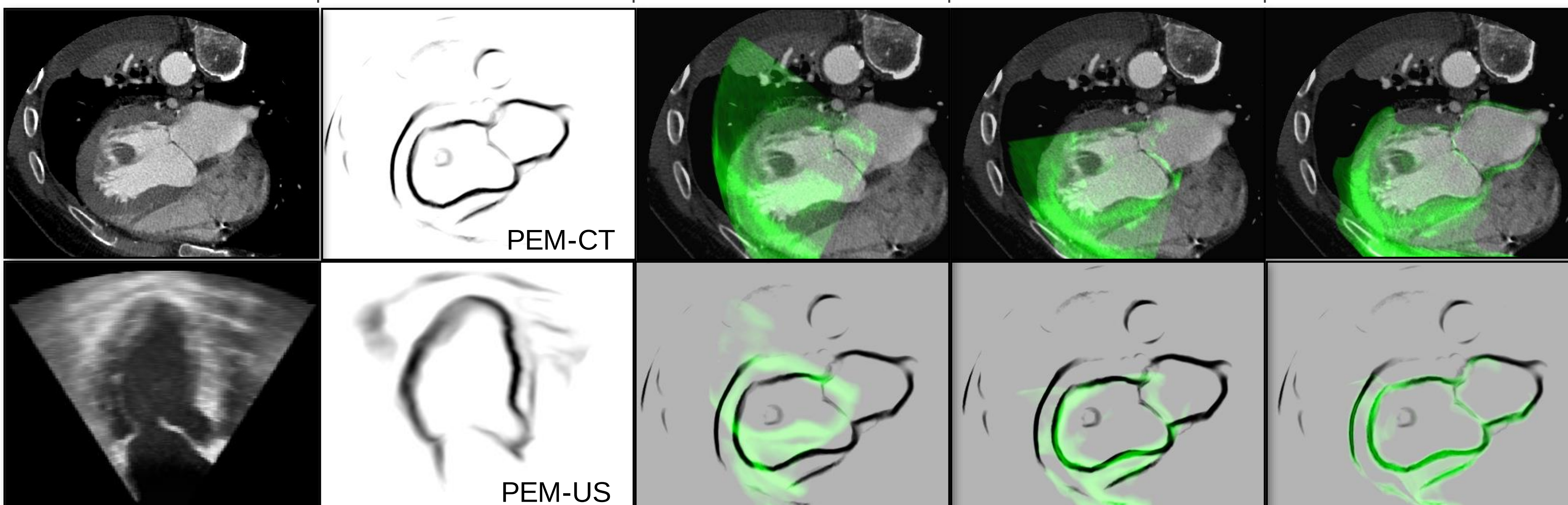
- Each voxel is voted for N_t x $(M_e)^3$
- N_t is the number of trees.
- All the votes are aggregated by averaging.

Dollar et al.: "Structured forests for fast edge detection." ICCV 2013

Proposed Multi-Modal Registration Framework

BioMedIA
Imperial College London

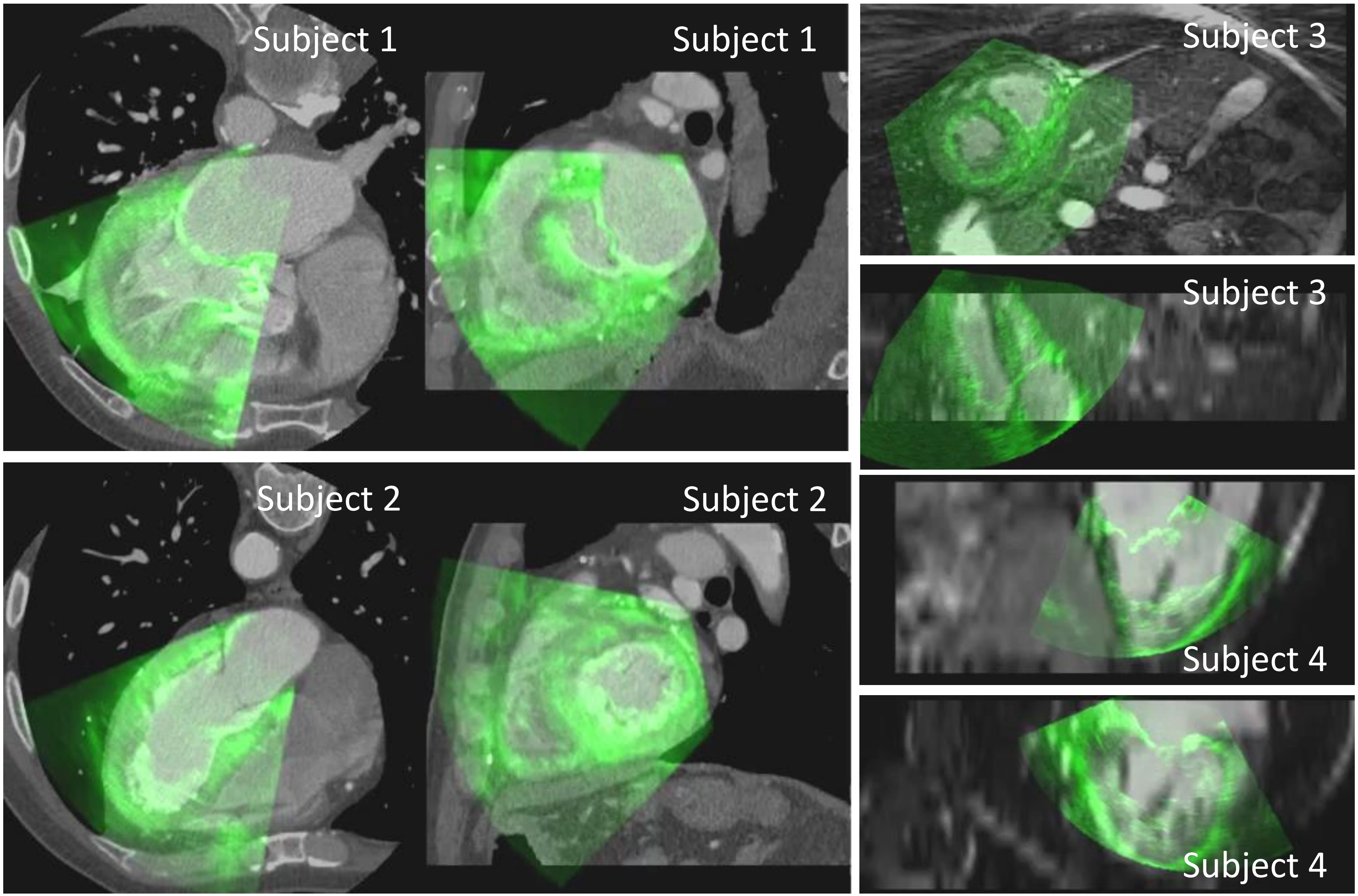
Input cardiac images	PEM representation	Initial Alignment of the images	Global alignment with robust block matching [2]	B-spline FFD based non-rigid registration [1]



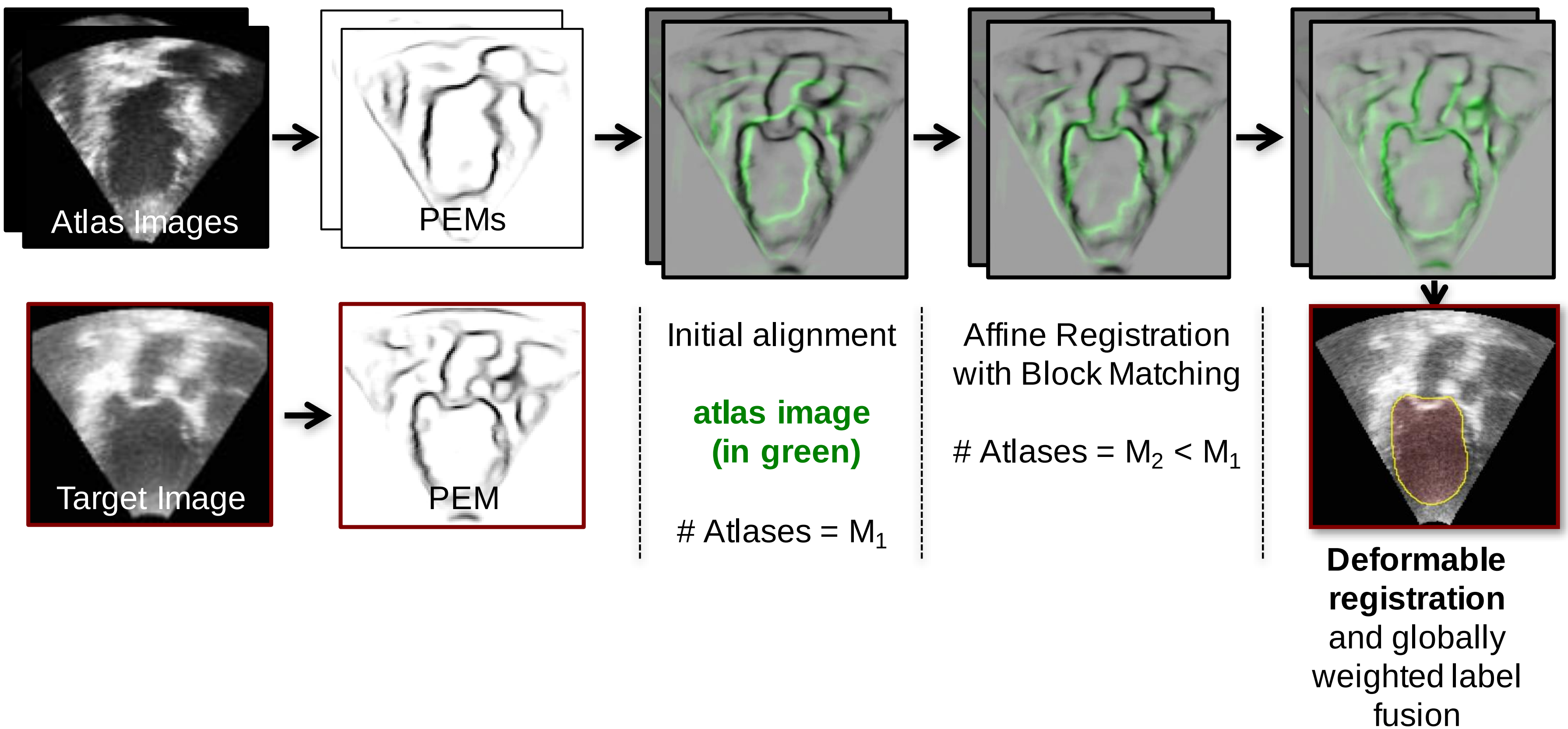
Computation Time (Quad-core 3.0GHz)	~20s per image		~21s per image	~73s per image

Rueckert et al.: "Non-rigid registration using free-form deformations: Application to breast MR images." TMI'99
Ourselin et al.: "Reconstructing a 3D structure from serial histological sections." Image and Vision Computing '01

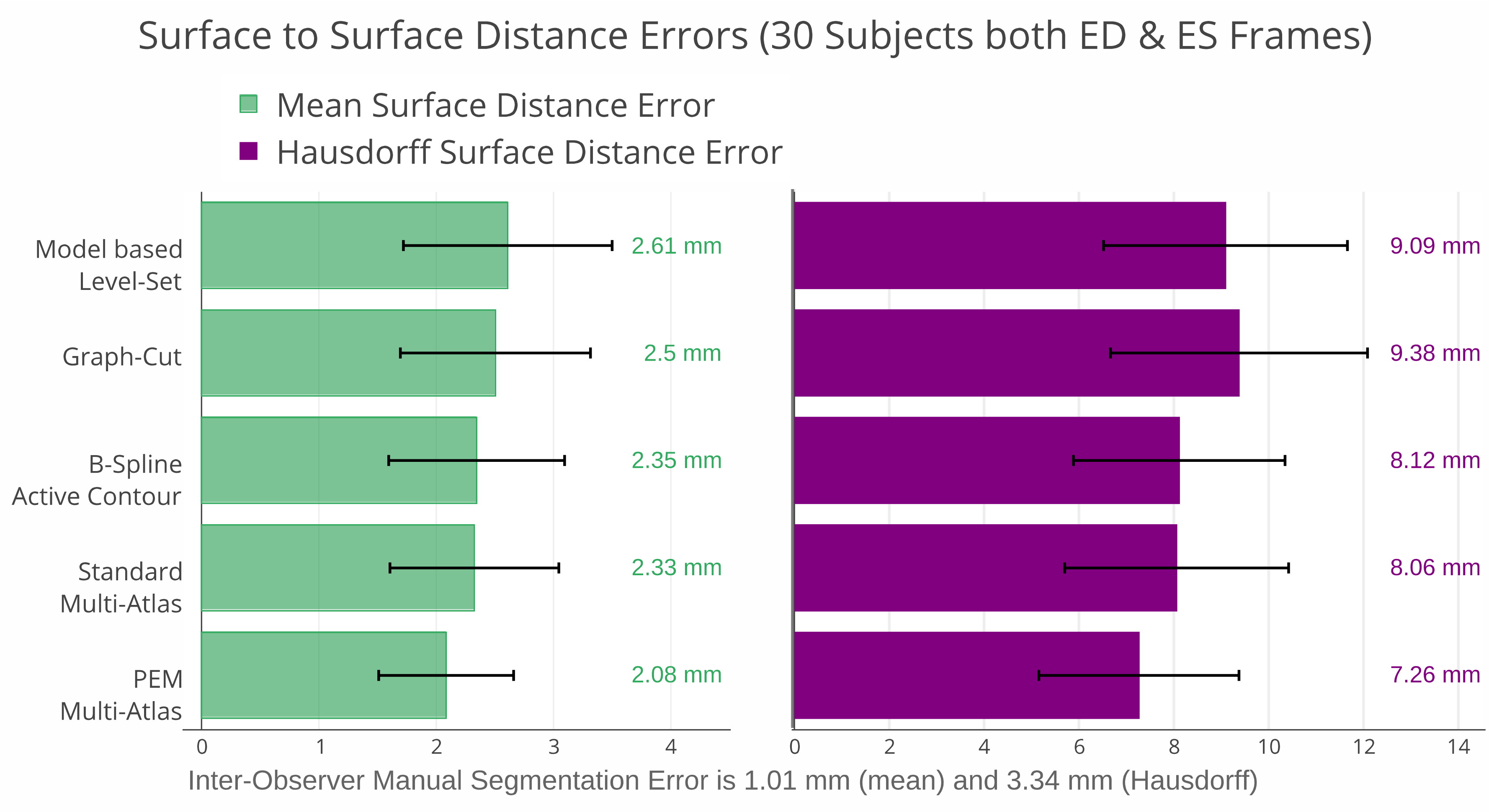
US/CT & US/MR Image Alignment



3DUS Image Registration with PEMs



Endocardial Surface Distance Errors



Imperial College London

Stratified Decision Forests for Accurate Anatomical Landmark Localization

Ozan Oktay, Wenjia Bai, Ricardo Guerrero, Martin Rajchl, Antonio de Marvao, Declan Regan, Stuart Cook, Mattias Heinrich, Ben Glocker, and Daniel Rueckert

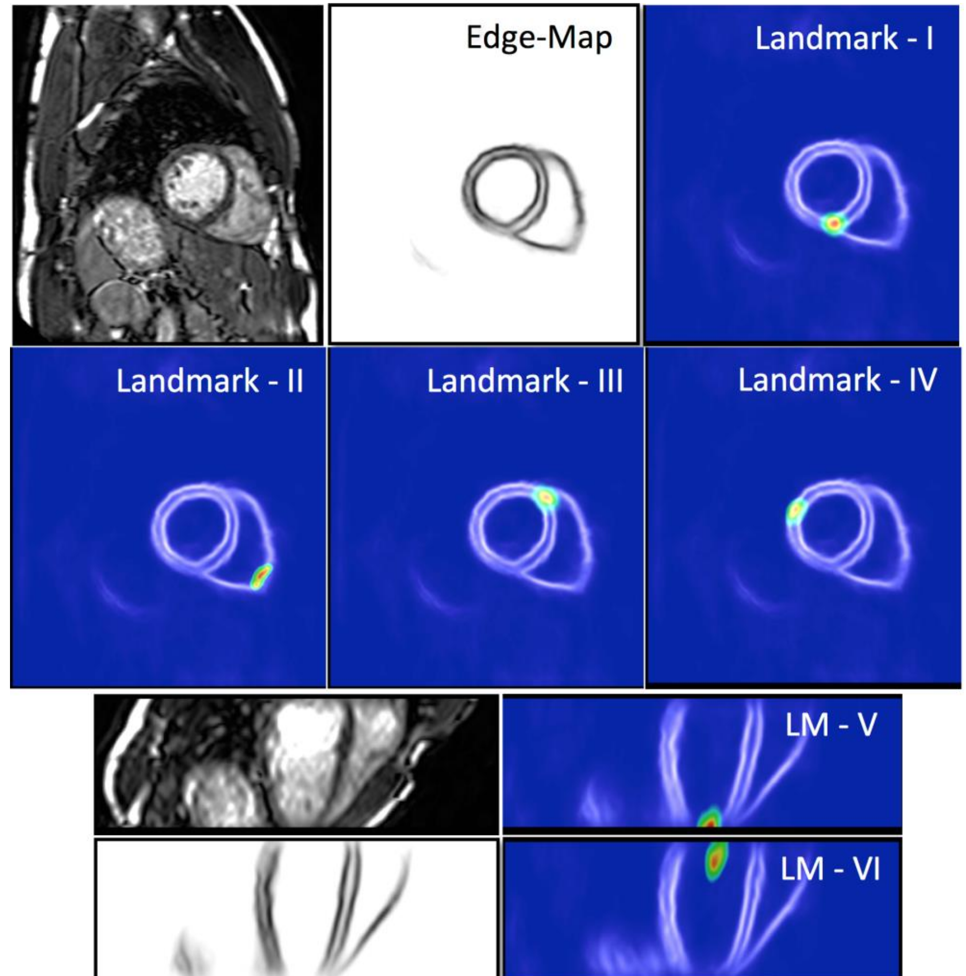
IEEE TMI, September 2016

Structured Regression Forest

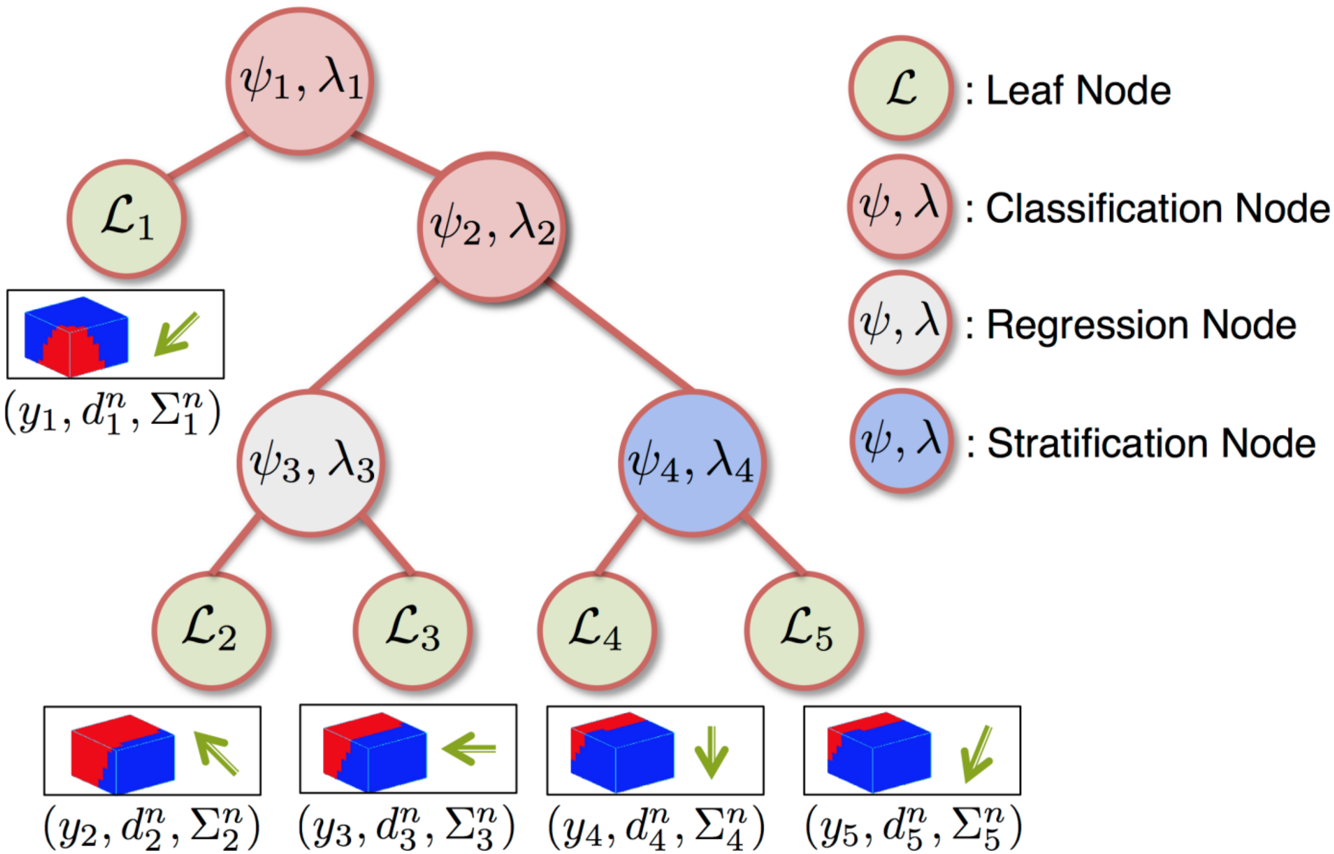
I. Anatomical landmark localization

II. Extracted boundaries regress the location of each landmark point

III. Similar approaches can be easily formulated with CNN models (e.g. YOLO v2)



Structured Regression Forest



Gall J., et al. "Class-Specific Hough Forests for Object Detection." CVPR 2009.
Criminisi A., et al. "Regression Forests for Efficient Anatomy Detection and Localization in CT Studies." MCV 2010.

Imperial College London

Attention Gated Networks: Learning to Leverage Salient Regions in Medical Images

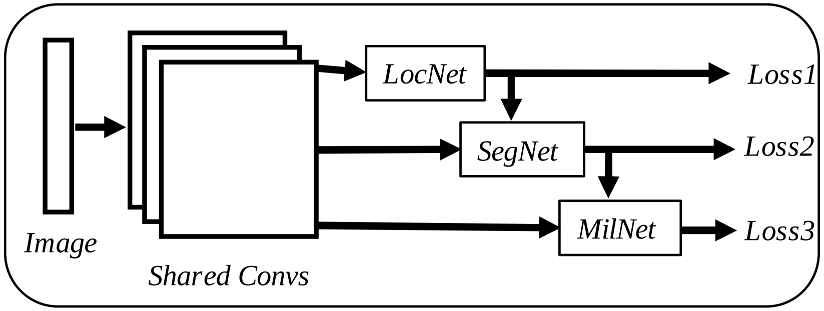
Ozan Oktay, Jo Schlemper, Michiel Schaap, Mattias Heinrich, Bernhard Kainz,
Ben Glocker, Daniel Rueckert

Medical Image Analysis Journal, Jan 2019

Cascaded Models in Image Analysis

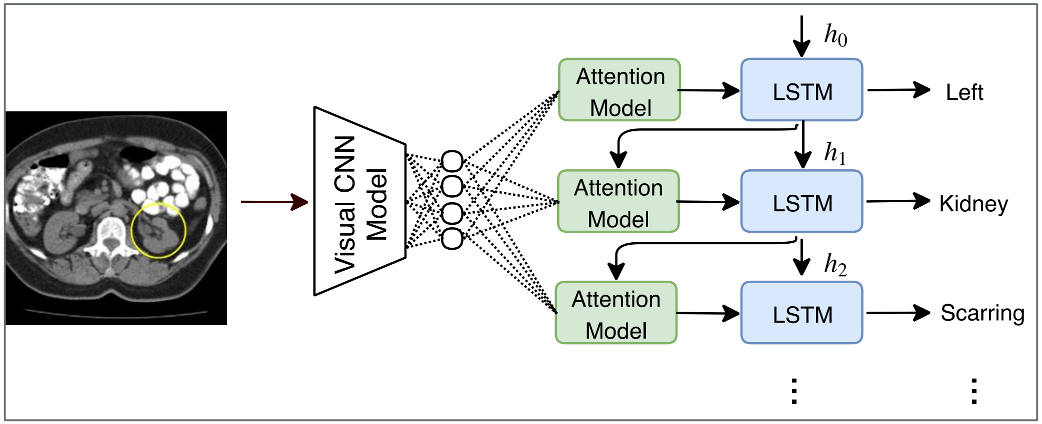
Cascaded models:

- Strategy: First localise then classify.
- GPU memory constraints.
- Solving simpler problems.
- Additional context information from preceding models.

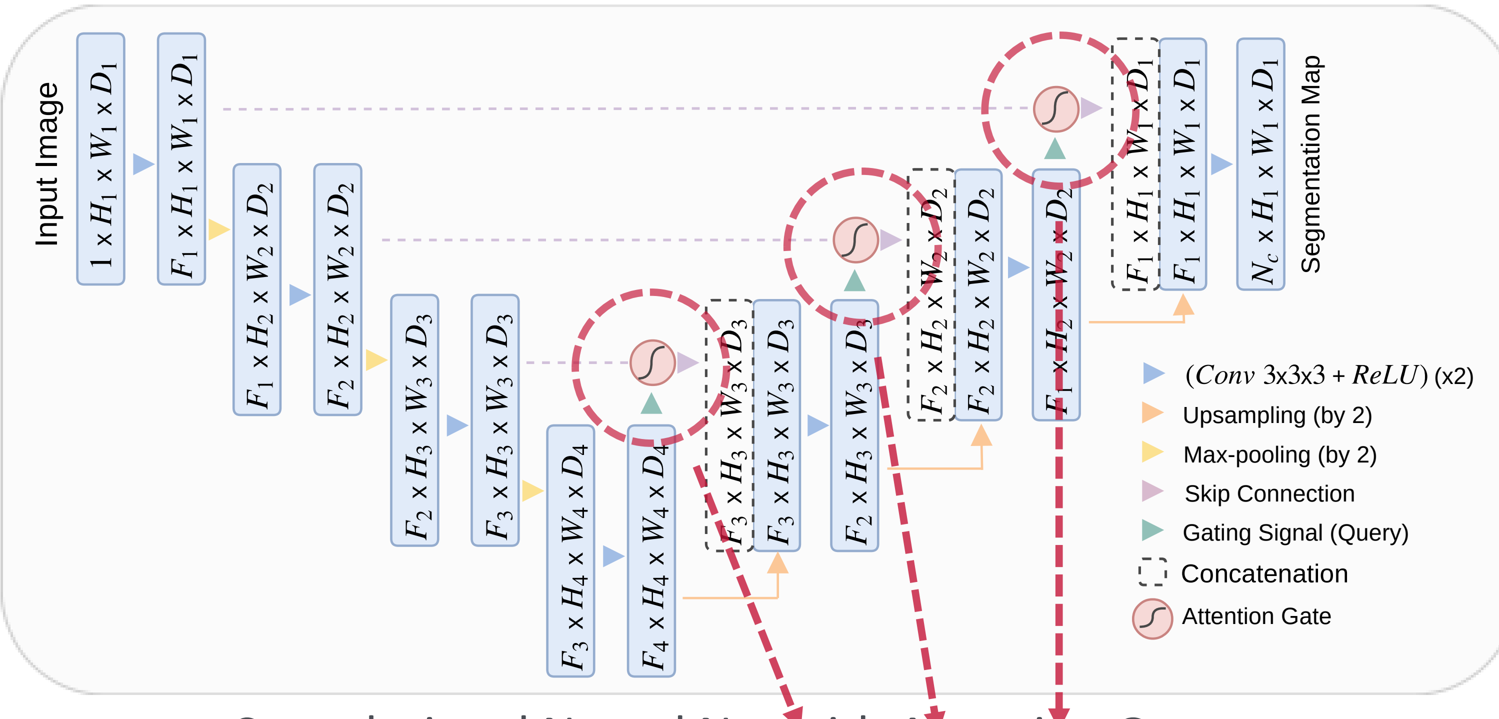


Potential Drawbacks:

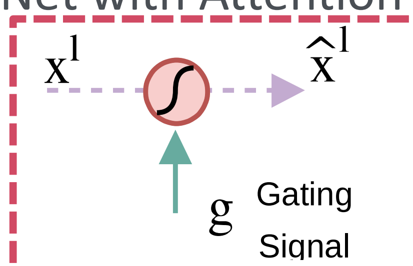
- Parameter & computation redundancy
- Multiple training schemes might be required



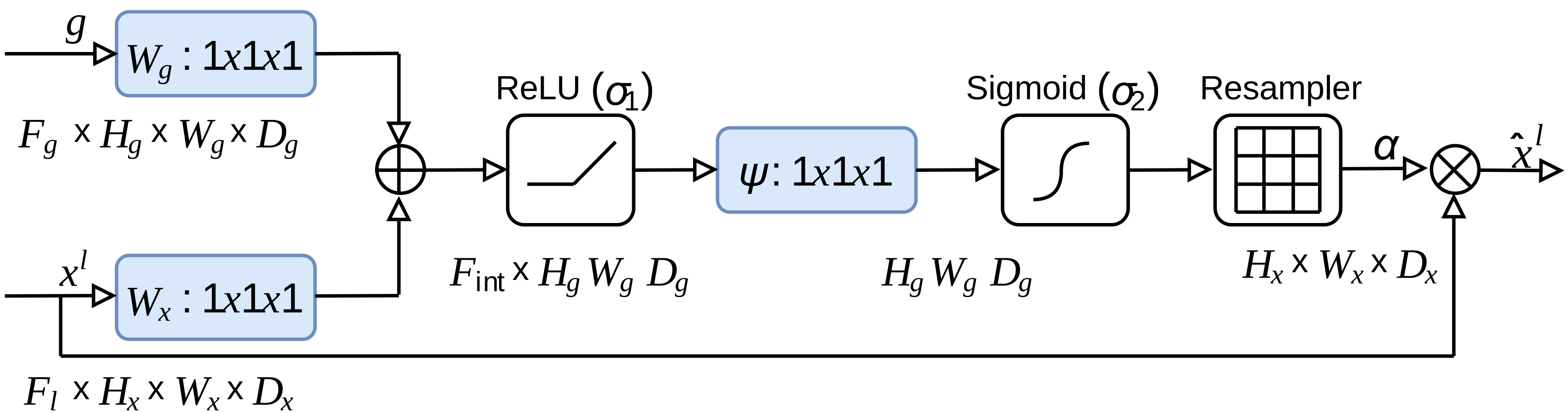
Attention Gates in CNN Models



Convolutional Neural Net with Attention Gates



Proposed Soft-Attention Gates



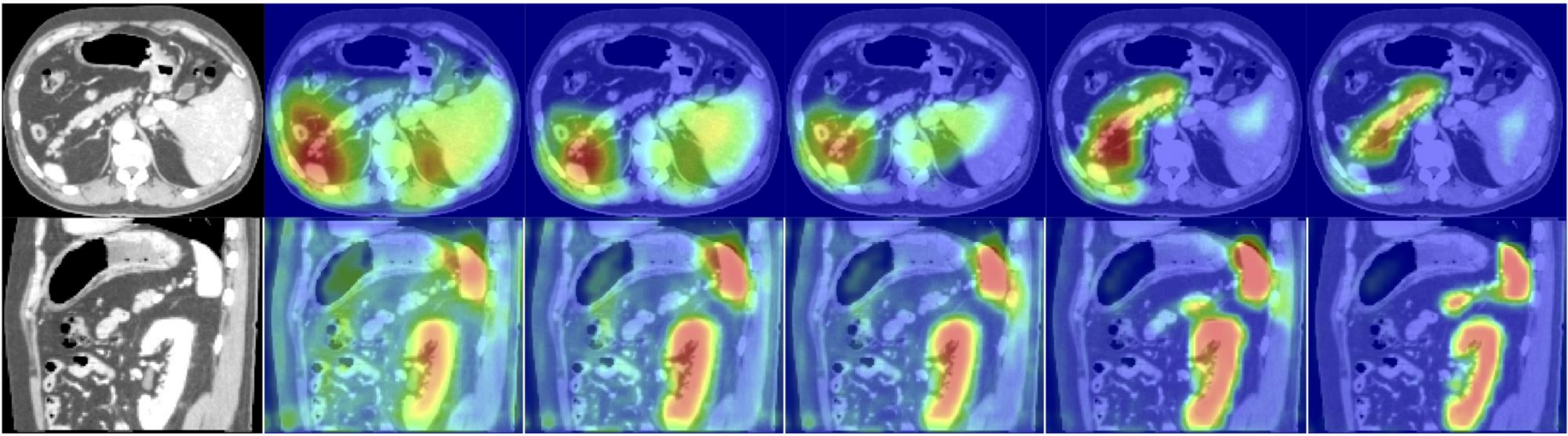
Concatenation (additive) based attention function

Final activation function has an influence on the training behaviour

$$\alpha = \sigma_2\left(\psi^T \left(\sigma_1\left(W_x^T x + W_g^T g + b_g \right) \right) + b_\psi \right)$$

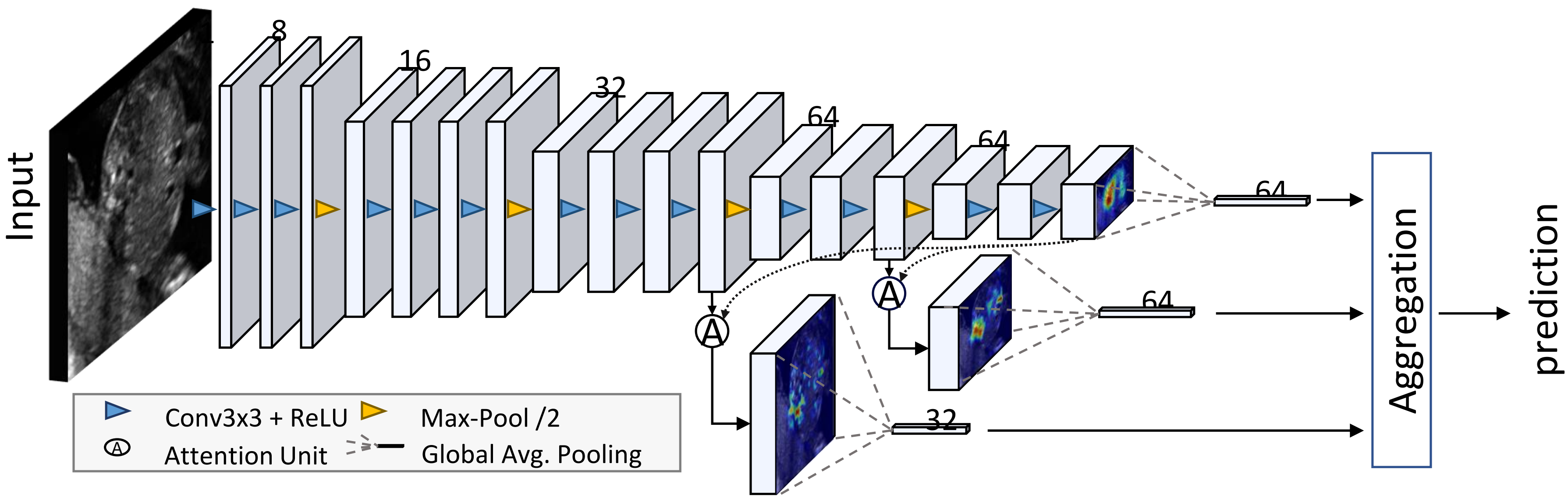
$$att(x, g) = \begin{cases} x^T g, & \text{dot} \\ x^T W_c\, g, & \text{general} \\ \psi^T tanh(W_c\, [x; g]), & \text{concat} \end{cases}$$

Attention Coefficients Across Different Training Epochs

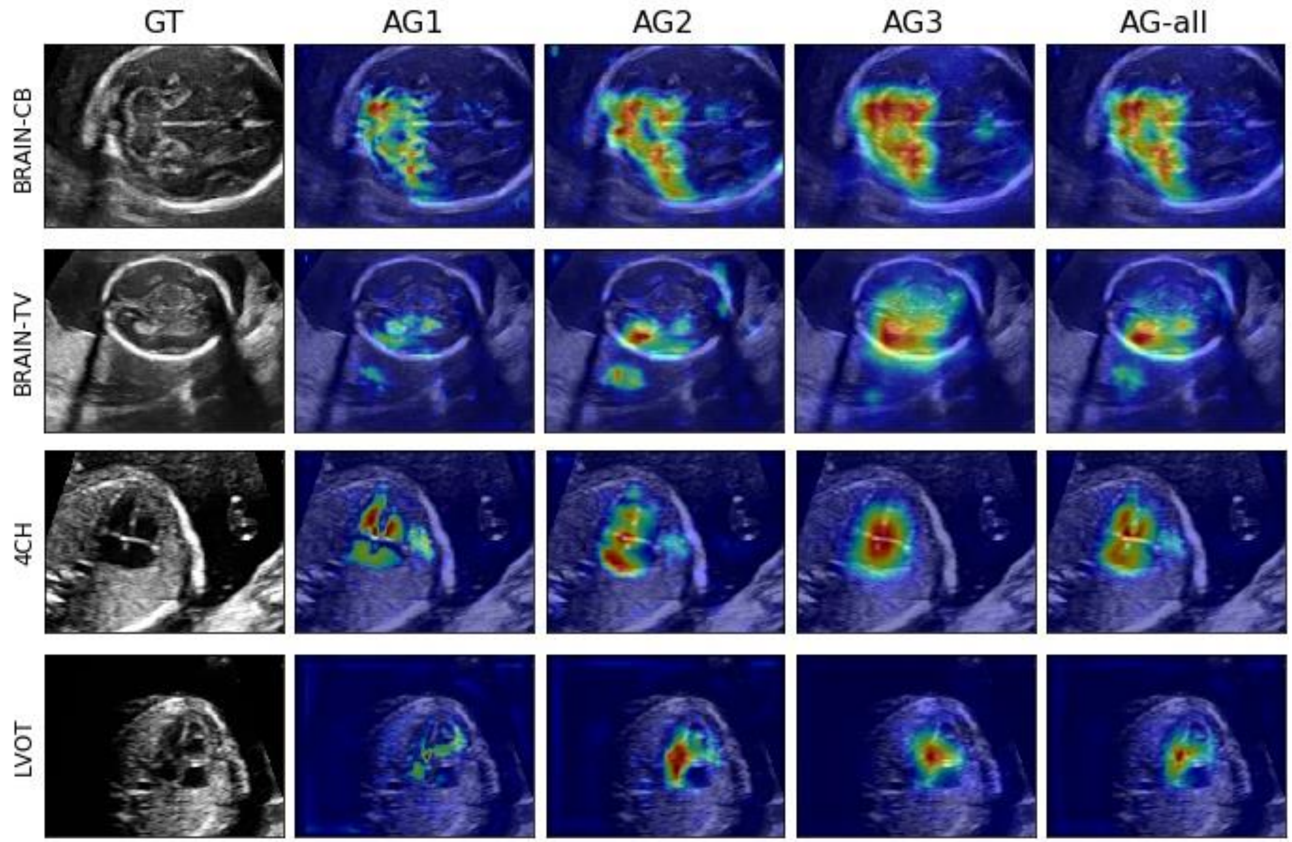


Attention coefficients across different training epochs
(kidneys, spleen, pancreas)

Attention Gates in Image Classification



Attention Maps at Different Scales



Adaptive pooling of feature maps with attention gates instead of using global aggregation

Acknowledgments

Thank you, Questions?

EPSRC
Engineering and Physical Sciences Research Council

Imperial College London