Machine Learning in Cardiac Imaging

Andrew King
School of Biomedical Engineering and Imaging Sciences
King's College London

smartheart

EPSRC
Pioneering research and skills

KING'S College LONDON

Outline

LEARNING FROM CLINICIANS

1. Incorporating non-imaging data into image-based machine learning

2. Exploiting multi-modal imaging data

MACHINE LEARNING IN ACQUISITION/RECONSTRUCTION/ANALYSIS

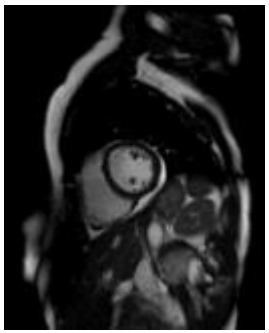
3. Automated quality control in large-scale imaging databases

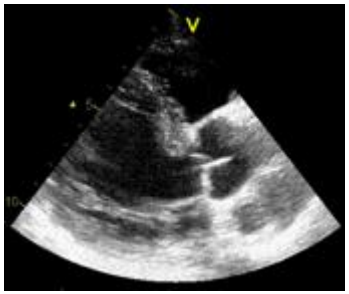
4. Machine learning for robust MR reconstruction

Learning from cardiologists

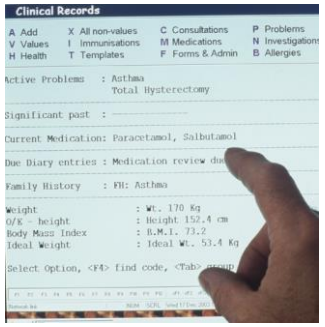
How do doctors make clinical decisions?

Imaging data





Patient record



Experience

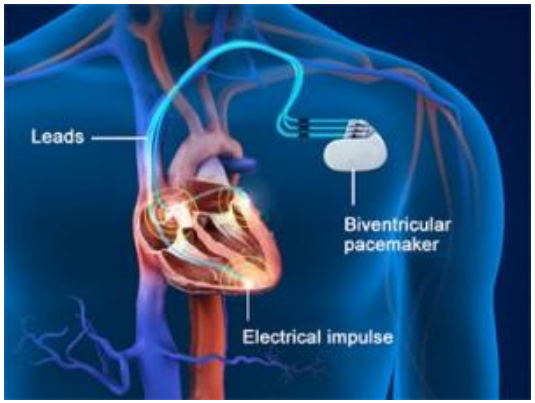
A. King – Machine Learning in Cardiac Imaging

KING'S College LONDON

1. Incorporating non-imaging data into image-based machine learning

Background

- Cardiac resynchronisation therapy (CRT) involves implanting a pacemaker to treat heart failure
- Using standard clinical selection criteria, ~30% of patients do not respond to treatment
- Research in the clinical literature has identified specific activation patterns that are associated with CRT response*, but these require manual inspection of imaging data by expert cardiologists



Aim

- Use machine learning to automatically learn imaging/non-imaging features to predict positive CRT response

* Sohal et al., JACC, 2013; Jackson et al, Heart Rhythm 2014

Data

Database:

34 patients selected for CRT

Pre-treatment tagged/cine/LGE MR

Follow-up data (positive/negative response to treatment)

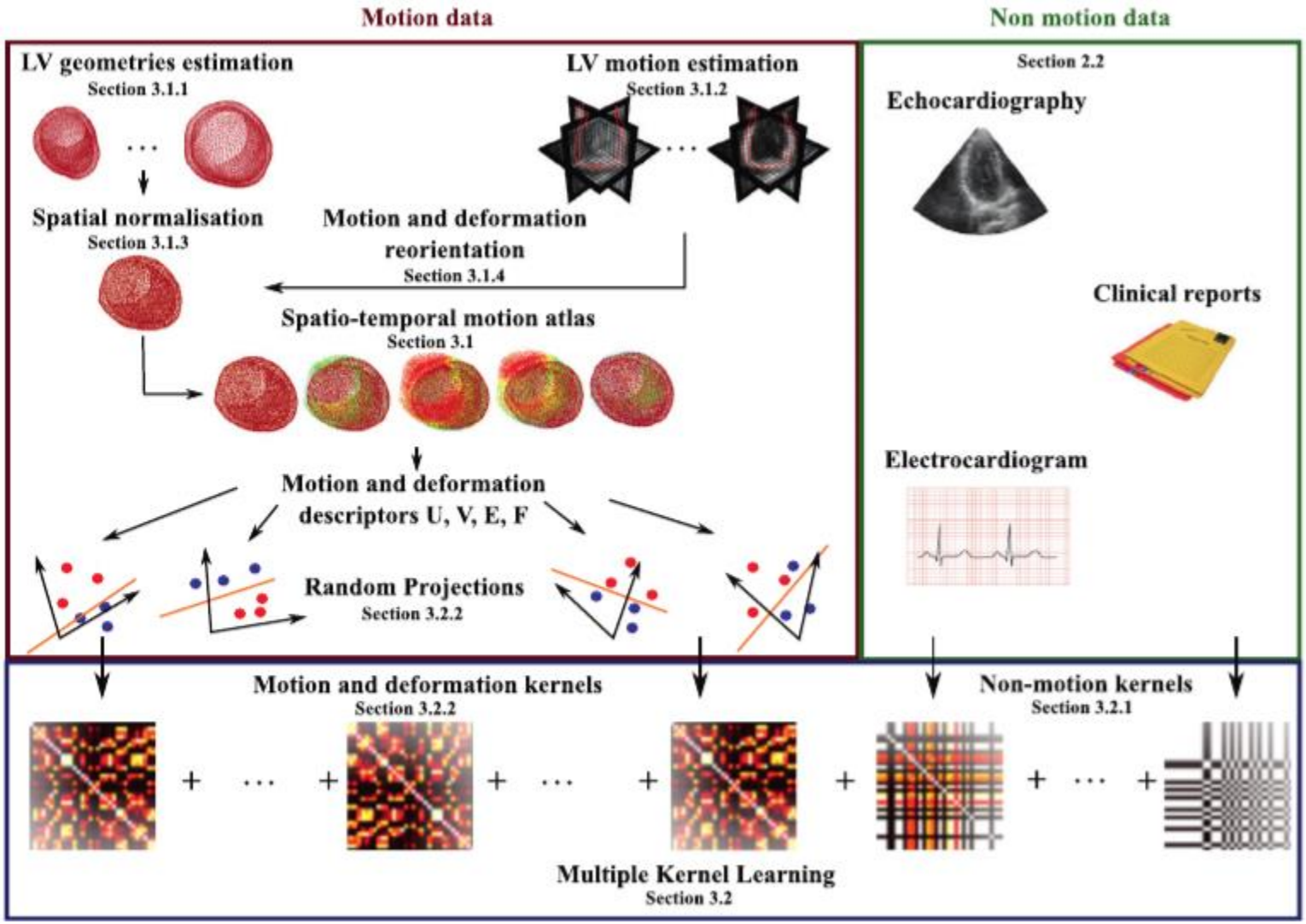
Non-imaging data:

Table 1
Description of the non-motion data derived from the clinical evaluation of the patient, from the ECG analysis and from 2D echocardiography imaging. For continuous data values, the third column reports mean and standard deviation over the entire cohort, while for categorical (binary) data values, the fourth column reports the counts of the corresponding categories.

Biomarker	Description	Mean/std dev	Frequency
Aetiology	Ischaemic/non-ischaemic	NA	13/21
EDV_m	End-Diastolic Volume from 3D geometry (cm^3)	281/127	NA
EDV	End-Diastolic Volume from 2D echo (cm^3)	214/91	NA
ESV	End-Systolic Volume from 2D echo (cm^3)	164/84	NA
EF	Ejection Fraction from 2D echo (%)	24.7/9.3	NA
Gender	Male/Female	NA	24/10
LBBB	Strict Left-Bundle Branch Block: yes/no	NA	23/11
NYHA	New York Heart Association classes (I-IV)	2.7/0.5	NA
QOL	Quality of Life questionnaire score	48/27	NA
QRS_d	QRS duration (ms)	146/22	NA
QRS_{cat}	QRS category $<$ 150 ms/ $>$ 150 ms	NA	18/16
Rhythm	Sinus/Atrial fibrillation	NA	28/6
6MWD	6 min walking distance (m)	269/137	NA

Peressutti et al, Med Image Anal, 2017

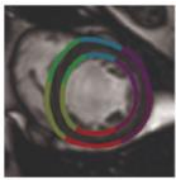
Multiple kernel learning



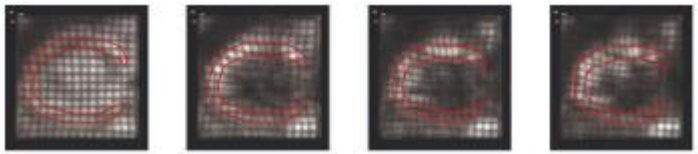
Peressutti et al, Med Image Anal, 2017

Motion atlas

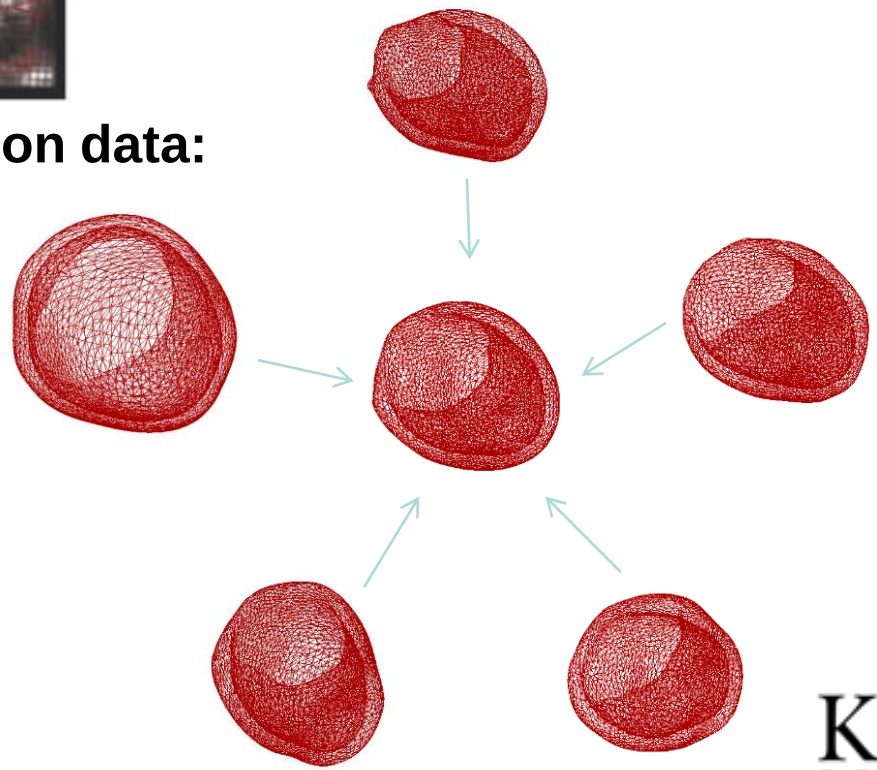
Segmentation of LV from cine MR:



Motion estimation from tagged MR:

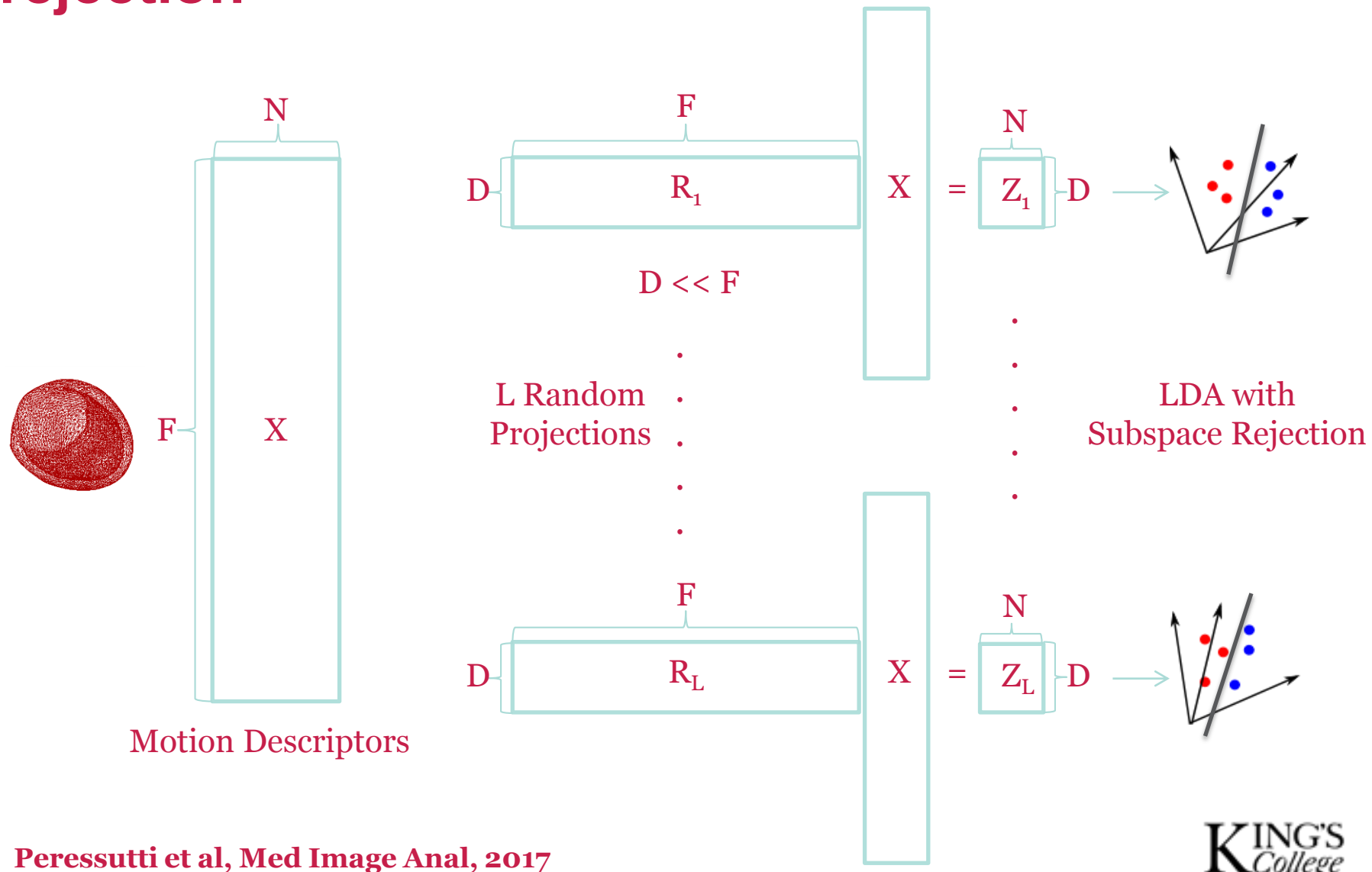


Transport of motion/deformation data:



Peressutti et al, Med Image Anal, 2017

Random projections with task-based subspace rejection



Peressutti et al, Med Image Anal, 2017

Multiple kernel learning for CRT response prediction - results

Classification results:

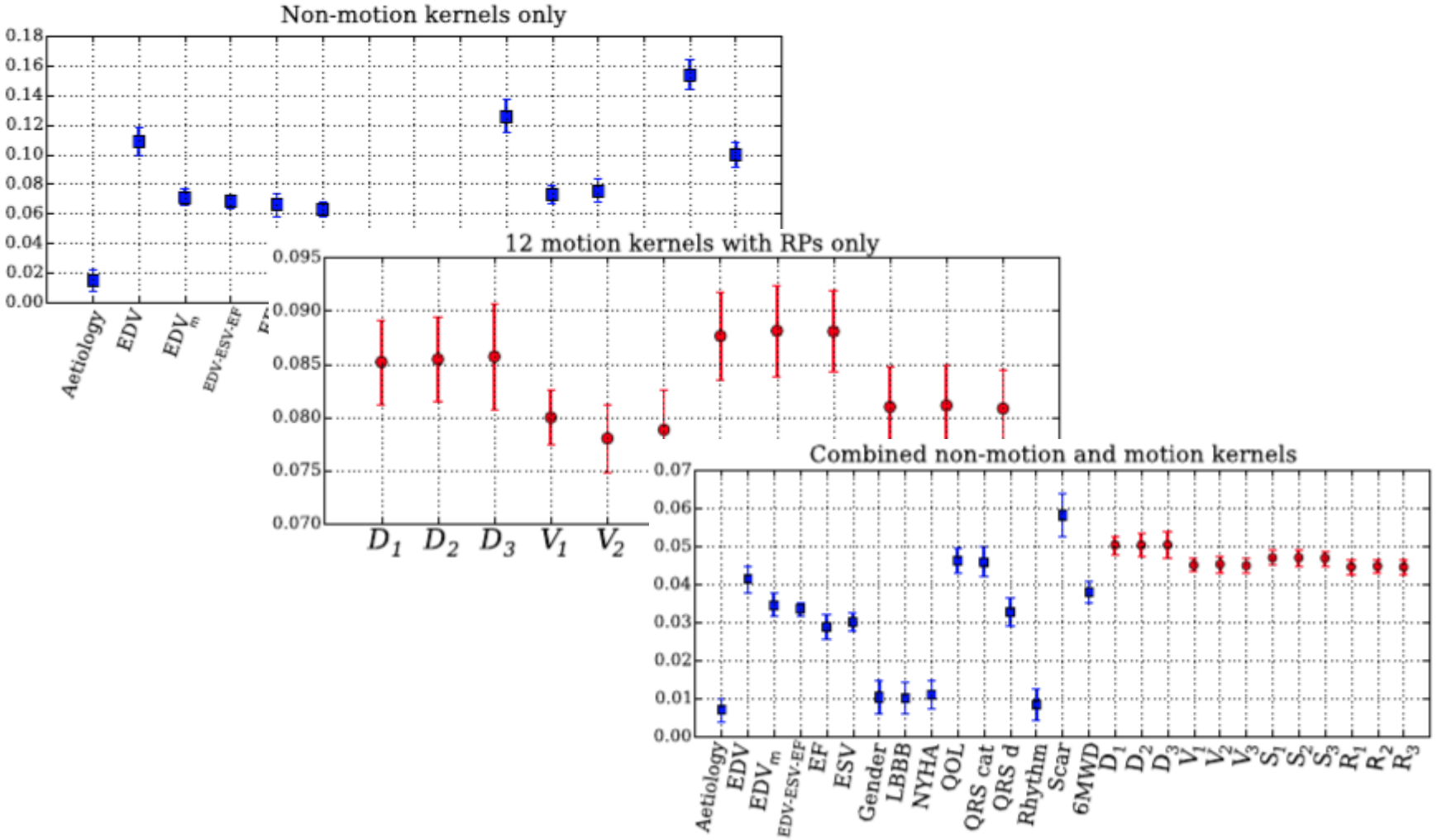
	Non-motion kernels only	Motion kernels only	Both motion and non-motion kernels
Accuracy	85.3	88.2	**91.2**
Sensitivity	100	100	**100**
Specificity	37.5	50	**62.5**
PPV	83.8	86.7	**89.7**
NPV	100	100	**100**

Sensitivity = proportion of responders chosen for treatment

Specificity = proportion of non-responders who would not be chosen for treatment

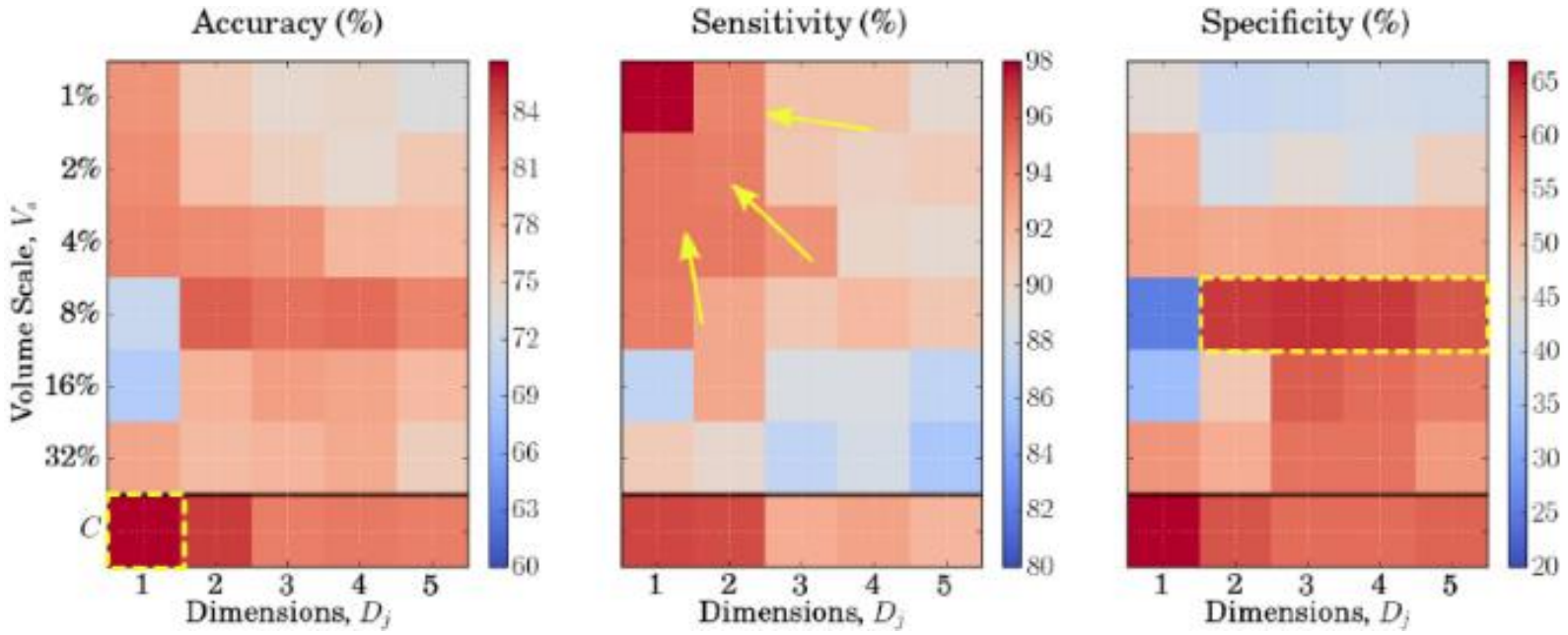
Peressutti et al, Med Image Anal, 2017

Multiple kernel learning for CRT response prediction – kernel weights



Peressutti et al, Med Image Anal, 2017

CRT response prediction – the role of spatial scale of the motion features



Sensitivity = proportion of responders chosen for treatment

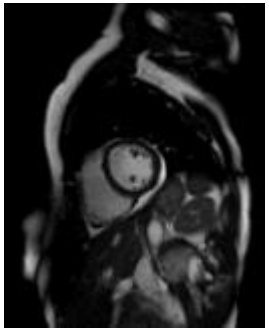
Specificity = proportion of non-responders who would not be chosen for treatment

Sinclair et al, Med Image Anal, 2018

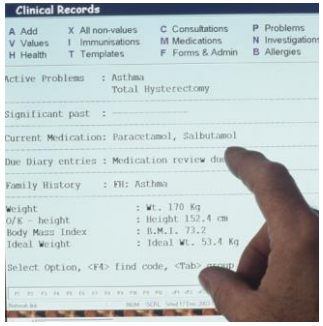
Learning from cardiologists

How do doctors make clinical decisions?

Imaging data



Patient record



Experience

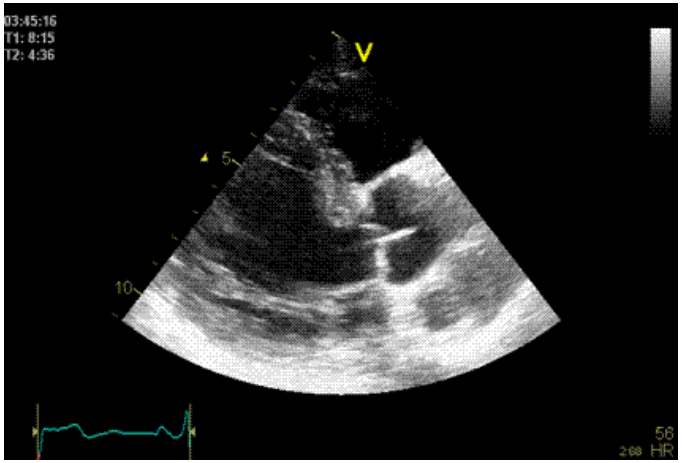
A. King – Machine Learning in Cardiac Imaging

KING'S College LONDON

2. Exploiting multi-modal imaging data

Background:

- MR is considered to be the ‘gold standard’ for cardiac functional assessment
- US is more commonly used due to its low cost, ease of use and portability



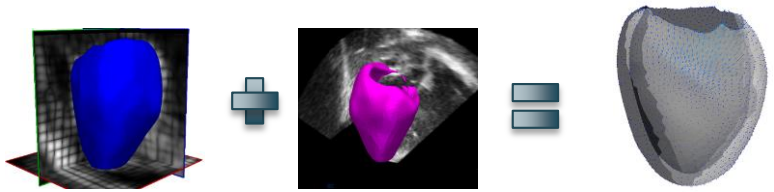
KING'S College LONDON

Exploiting multi-modal imaging data

Aim:

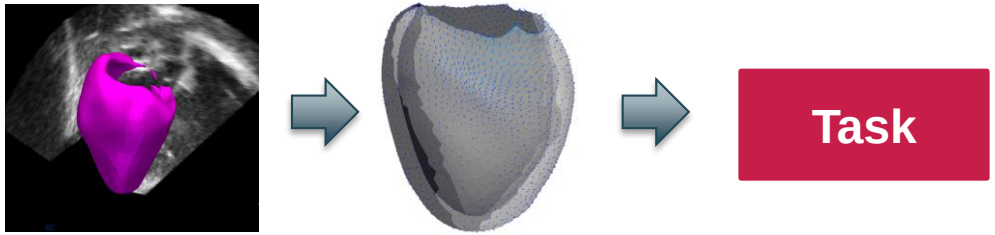
- Use a database of paired MR/US data sets to exploit multimodal data in a diagnostic pipeline
- Pipeline should be based on data from a single modality (US)

Train:



MR US Multimodal motion atlas

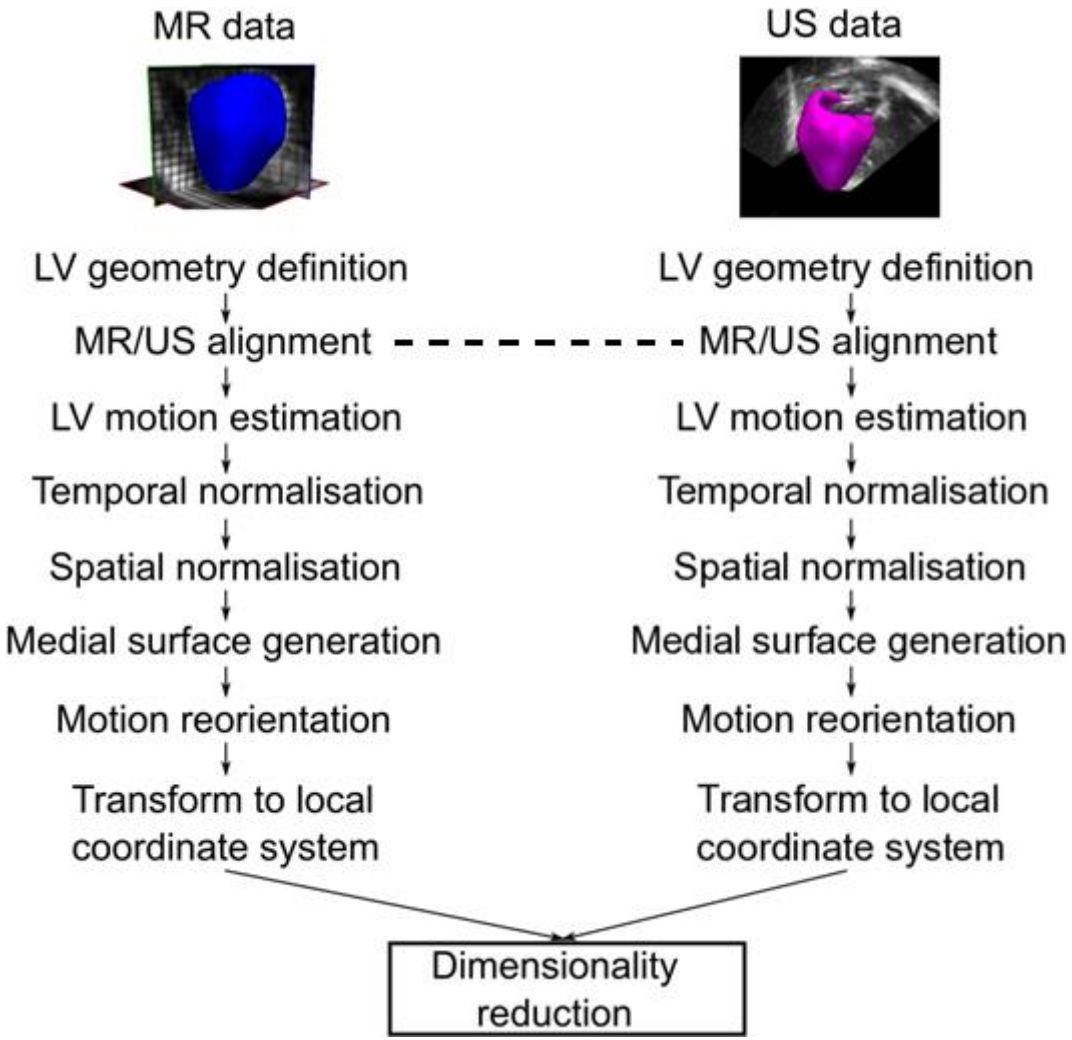
Test:



US Multimodal motion atlas

KING'S College LONDON

A multimodal cardiac motion atlas



Puyol-Anton et al., Med Image Anal, 2017

Multiview dimensionality reduction

Multi-view dimensionality reduction algorithms

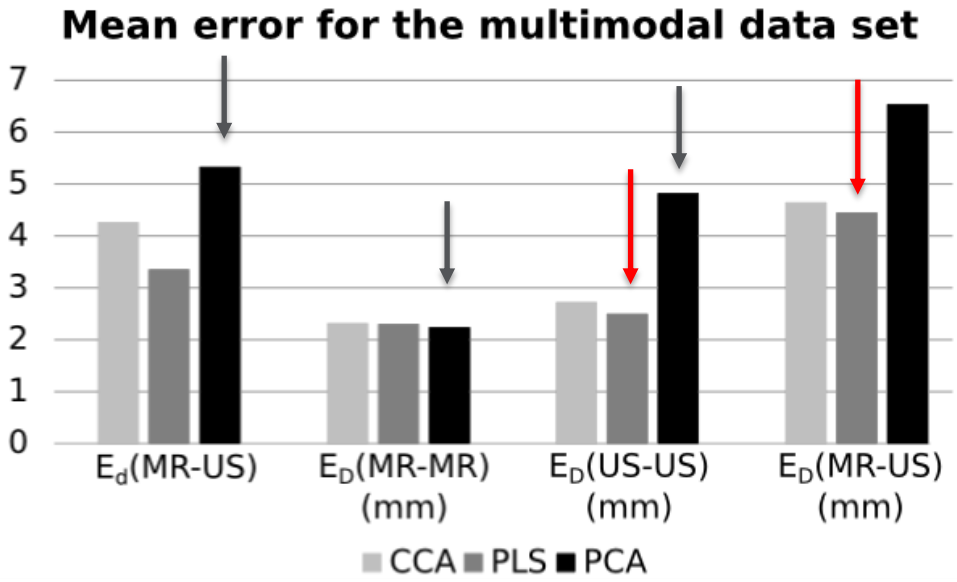
- *Canonical correlation analysis (CCA)* on MR/US displacements
  - Determines directions where X and Y have maximum *correlation* between modalities
- *Partial least squares regression (PLS)* on MR/US displacements
  - Determines directions where X and Y have maximum *covariance* between modalities

Single view dimensionality reduction algorithm

- *Principal component analysis (PCA)* on MR displacements only
  - Determines directions of maximum data *variance* in single modality

Puyol-Anton et al., Med Image Anal, 2017

Multimodal atlas: results



- CCA and PLS have lower errors than PCA.

- PLS has lower errors and better reconstructed volumes compared to CCA.

E_d(MR-US) – embedding error
E_D(MR-MR) – Reconstruction error for MR
E_D(US-US) – Reconstruction error for US
E_D(MR-US) – Prediction error

Puyol-Anton et al., Med Image Anal, 2017

The task

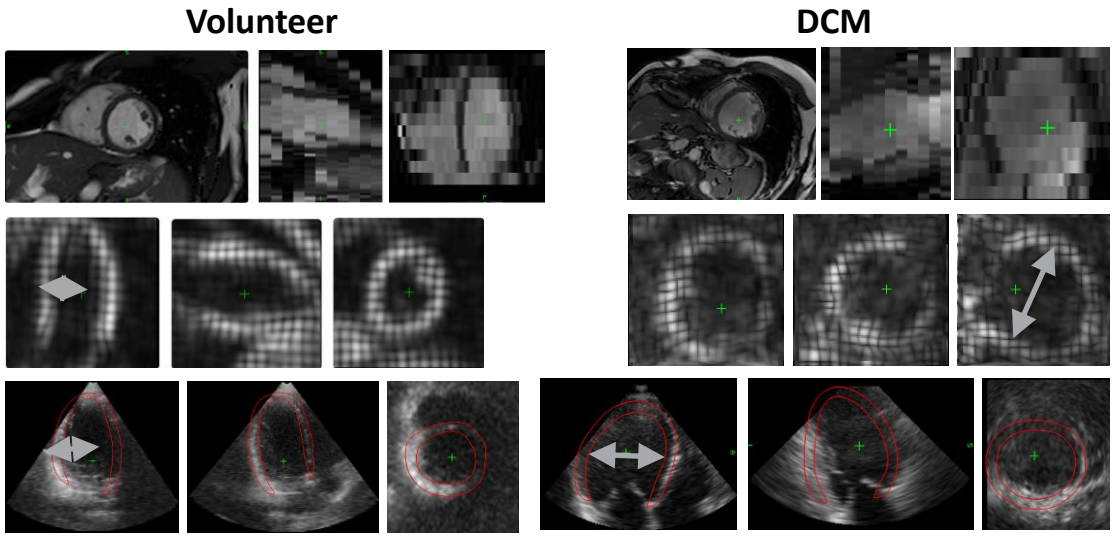
Classification of normal vs. dilated cardiomyopathy:

Clinical database:

- 50 healthy volunteers
- 14 dilated cardiomyopathy patients

Image protocol (LV):

- Multi-slice short-axis MR sequence
- 3D tagged MR sequence (3DTAG).
- 3D apical ultrasound sequence (3DUS).

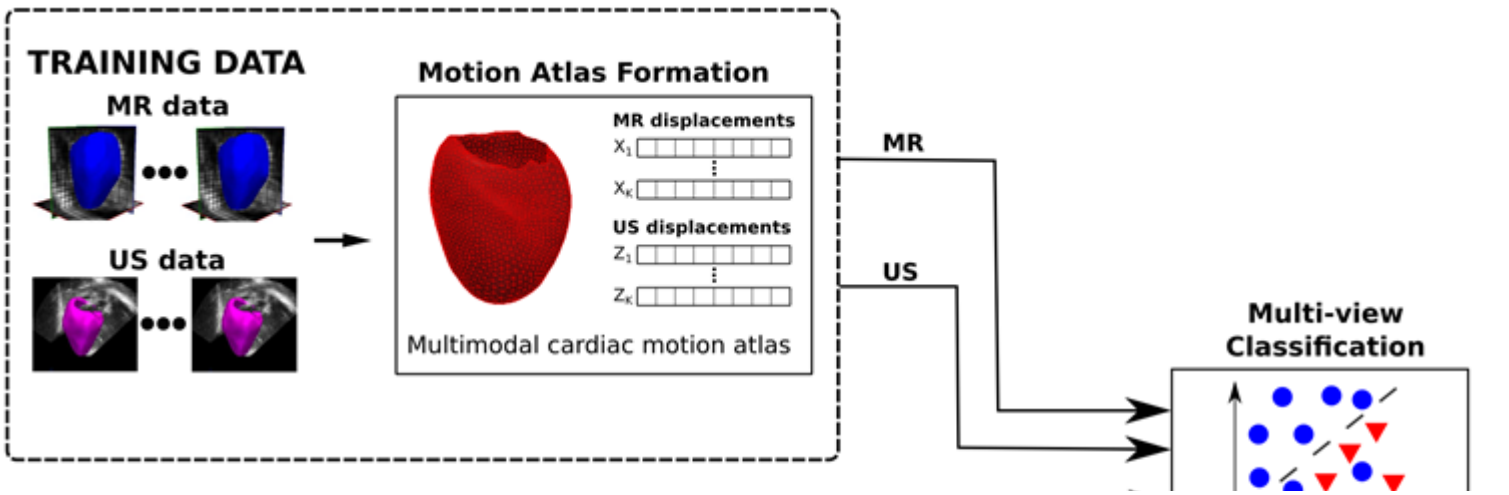


Stratified k-fold cross-validation

- Imbalanced data
- # low dimensions set to retain 90% of the variance
- 8 folds and 100 repetitions

Puyol-Anton et al., IEEE Trans Biomed Eng, 2018

Multiview classification



2.a. Single modality US method

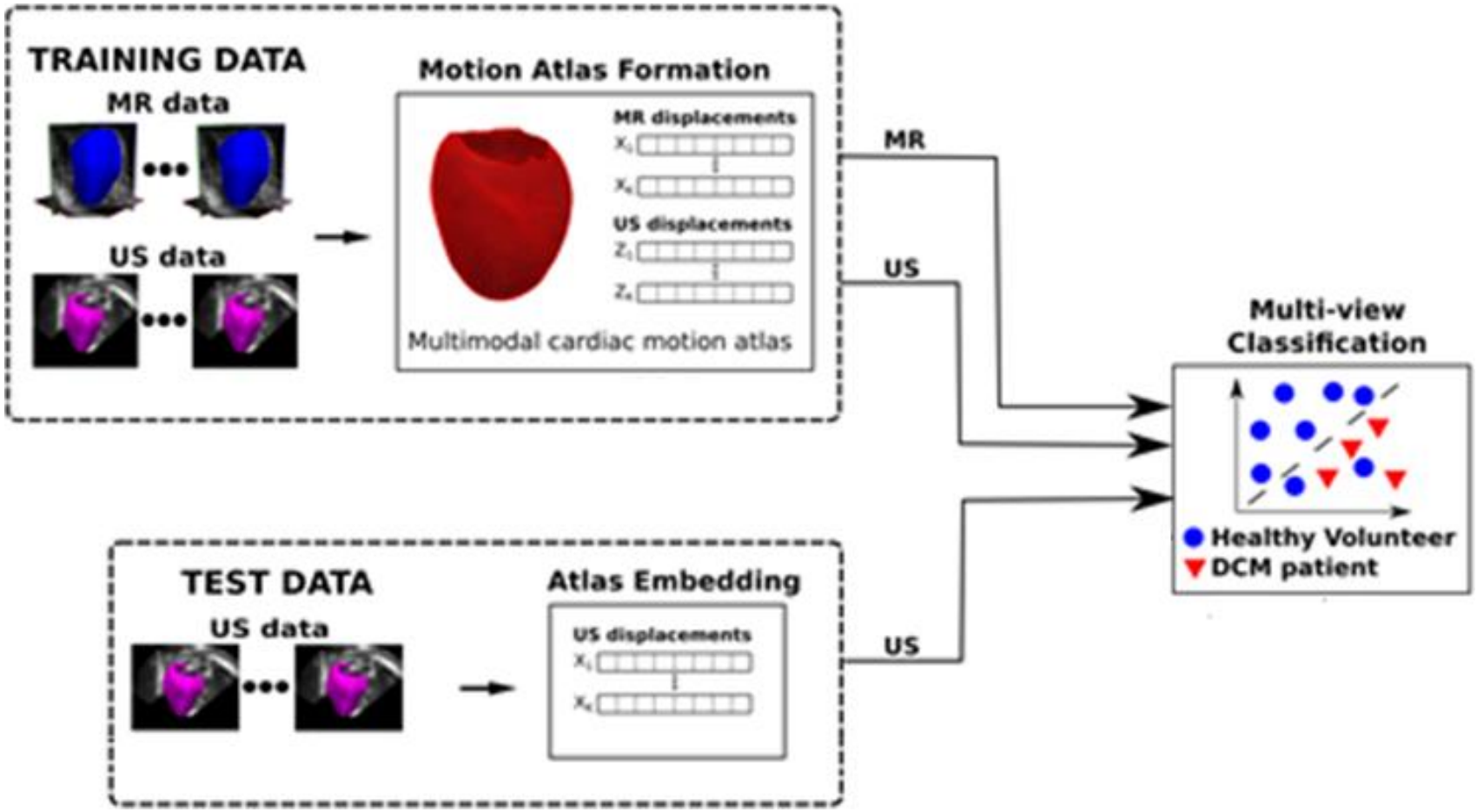
PCA_{US} + LDA	74.79 (15.8)	71.05 (28.6)	78.57 (10.1)
PCA_{US} + SVM_{rbf}	87.32 (12.9)	84.50 (23.2)	90.14 (6.6)

3. Multiple modality two-step method

PLS + LDA	80.07 (16.8)*	75.51 (27.1)	81.86 (9.5)
PLS + SVM_{rbf}	90.39 (12.1)*	87.50 (21.8)	90.57 (10.9)

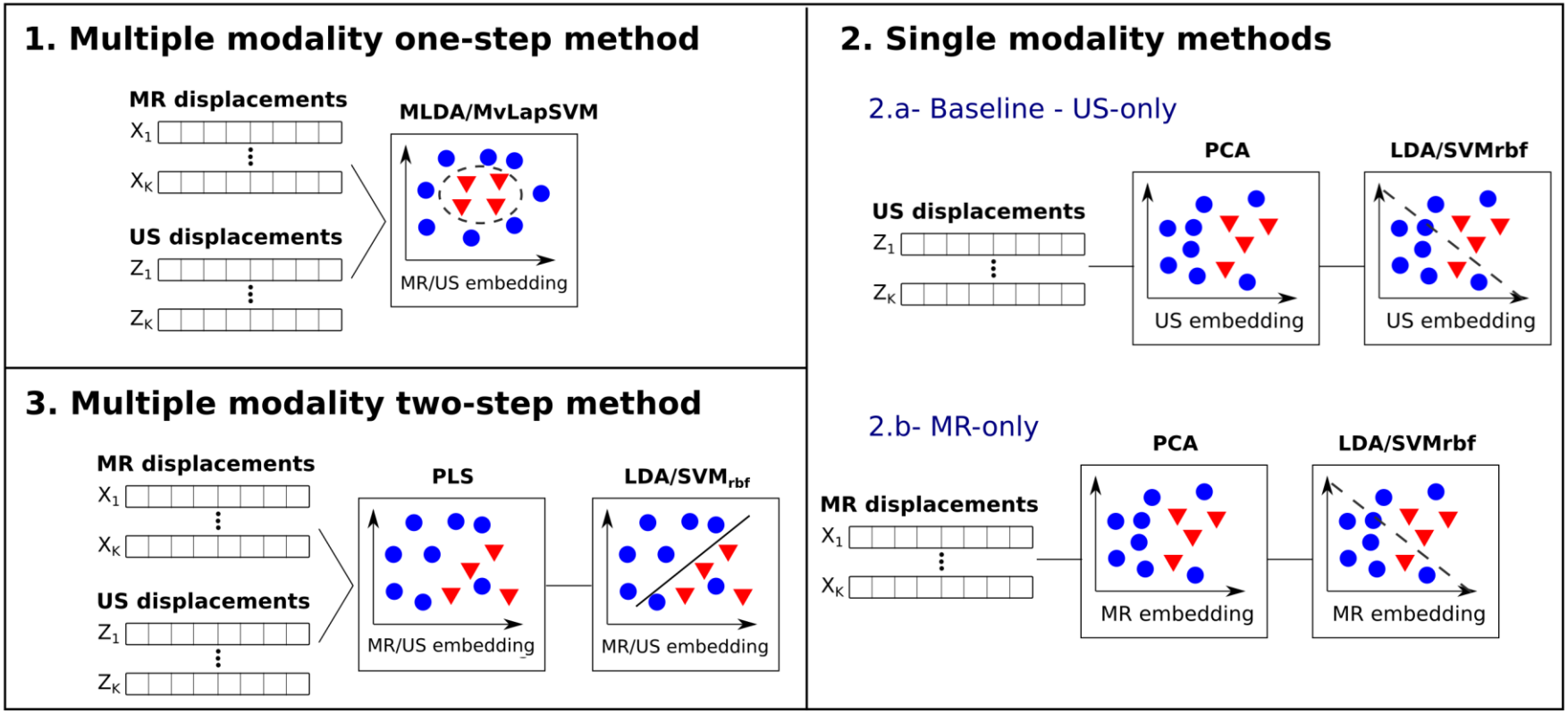
Puyol-Anton et al., IEEE Trans Biomed Eng, 2018

Multiview learning



Puyol-Anton et al., IEEE Trans Biomed Eng, 2018

Multiview learning



Puyol-Anton et al., IEEE Trans Biomed Eng, 2018

Multiview learning - results

Proposed method	BACC (%)	SEN (%)	SPE (%)
1. Multiple modality one-step method			
MLDA	**82.18 (15.0)***	80.50 (26.5)	83.86 (9.9)
MvLapSVM	**92.71 (10.4)***	89.00 (20.8)	95.14 (6.8)
Comparative approaches	**BACC (%)**	**SEN (%)**	**SPE (%)**
2.a. Single modality US method			
PCA_{US} + LDA	74.79 (15.8)	71.05 (28.6)	78.57 (10.1)
PCA_{US} + SVM_{rbf}	87.32 (12.9)	84.50 (23.2)	90.14 (6.6)
2.b. Single modality MR method:			
PCA_{MR} + LDA	84.21 (15.4)*	74.00 (28.8)	90.43 (6.8)
PCA_{MR} + SVM_{rbf}	90.89 (11.7)*	86.50 (22.3)	95.29 (6.7)
3. Multiple modality two-step method			
PLS + LDA	80.07 (16.8)*	75.51 (27.1)	81.86 (9.5)
PLS + SVM_{rbf}	90.39 (12.1)*	87.50 (21.8)	90.57 (10.9)

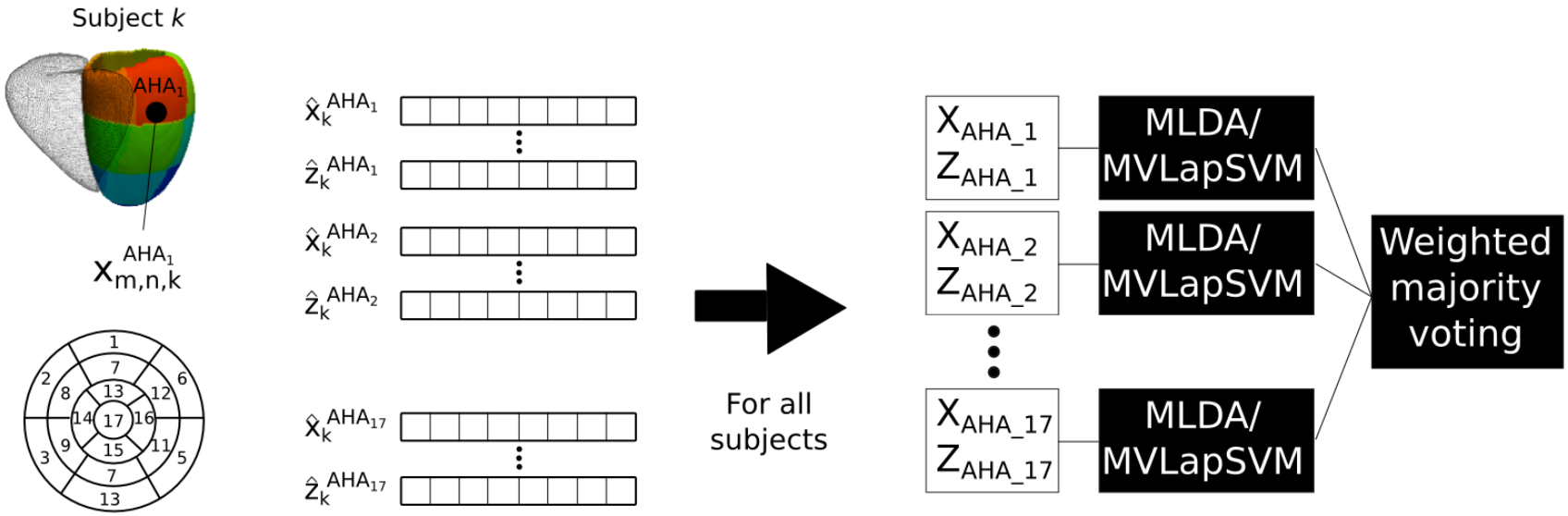
- One-step multiview learning outperforms two-step approach
- One-step multiview learning performs almost as well as the 'gold standard' of MR-based classification

BACC: balanced accuracy
SEN: sensitivity
SPE specificity

* Student's t-test (99% confidence)

Puyol-Anton et al., IEEE Trans Biomed Eng, 2018

Regional multiview learning



Weighted majority voting:

For each AHA segment each subject is classified. The results are combined using weighted majority voting with weights determined at training by a randomised search on hyper parameters

Puyol-Anton et al., IEEE Trans Biomed Eng, 2018

Regional multiview learning - results

Global Methods	BACC (%)	SEN (%)	SPE (%)
MLDA	82.18 (15.0)	80.50 (26.5)	83.86 (9.9)
MvLapSVM	92.71 (10.4)	89.00 (20.8)	95.14 (6.8)
Regional Methods	**BACC (%)**	**SEN (%)**	**SPE (%)**
MLDA	87.71 (12.6)*	85.00 (23.1)	90.43 (6.7)
MvLapSVM	94.32 (11.1)*	93.00 (17.5)	96.57 (6.2)

- Regional approach has higher accuracy
- Highest accuracy was 94% using regional MvLapSVM.

BACC: balanced accuracy
SEN: sensitivity
SPEL specificity

* Student's t-test (99% confidence)

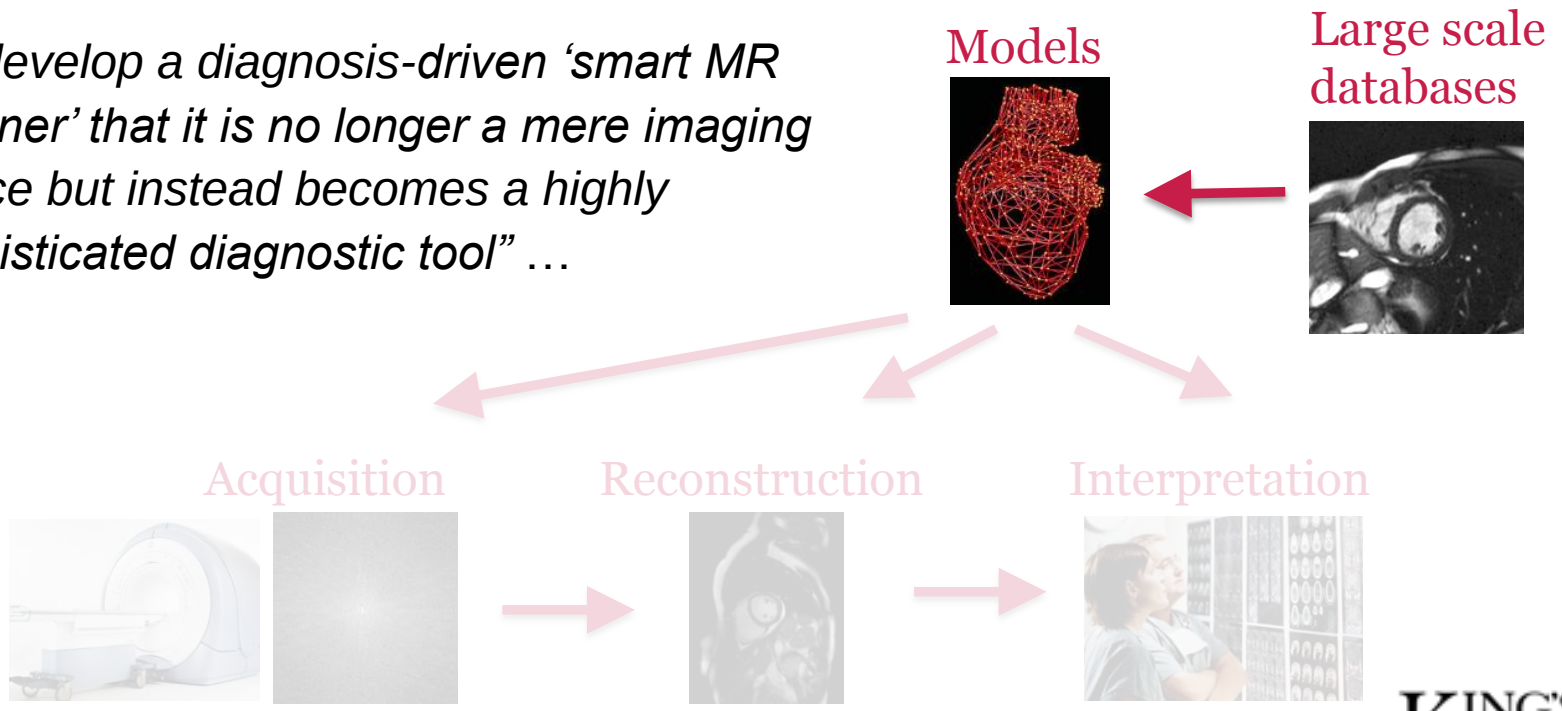
Puyol-Anton et al., IEEE Trans Biomed Eng, 2018

SmartHeart project

5 year grant involving:

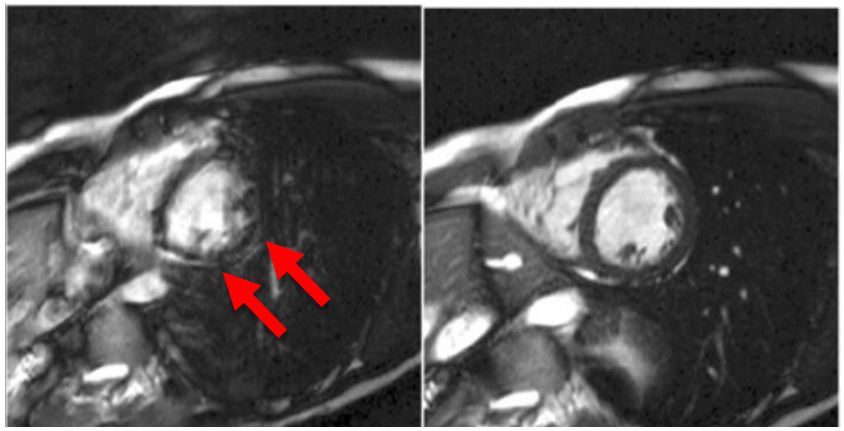
- King's College London
- Imperial College London
- Queen Mary University London
- Oxford University

"To develop a diagnosis-driven 'smart MR scanner' that it is no longer a mere imaging device but instead becomes a highly sophisticated diagnostic tool" …



3. Automated quality control in large-scale imaging databases

Cardiac MR



Adopted from Ferreira et al., JCMR, 2013.

- Need for high quality images
- Wide range of artefacts
- Manual labelling tedious for large datasets
- Need for automatic quality assessment tools

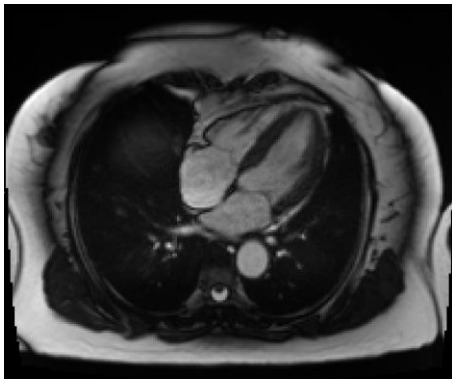
biobank uk
Improving the health of future generations

- UK Biobank is a large scale database of imaging/non-imaging data
- Will eventually consist of cardiac MR images from 100,000 subjects (currently ~27,000)

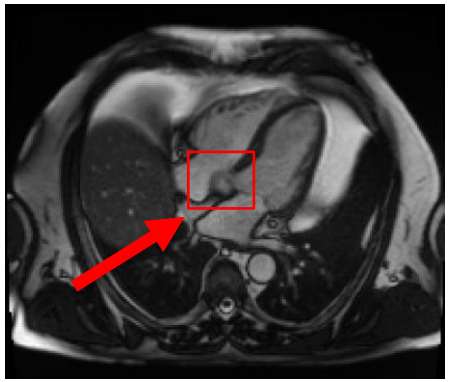
Cardiac MR quality issues

1. **Off-axis (4ch)***
   - Left Ventricular Outflow Tract
   - 5 chamber look

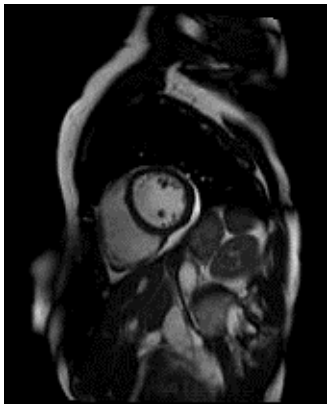
2. **Motion related artefacts (SAX)$**
   - Breathing
   - Mis-triggering
   - Arrhythmia



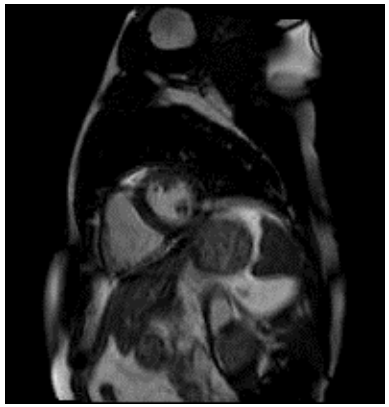
Good Planning



Bad Planning



Good Quality

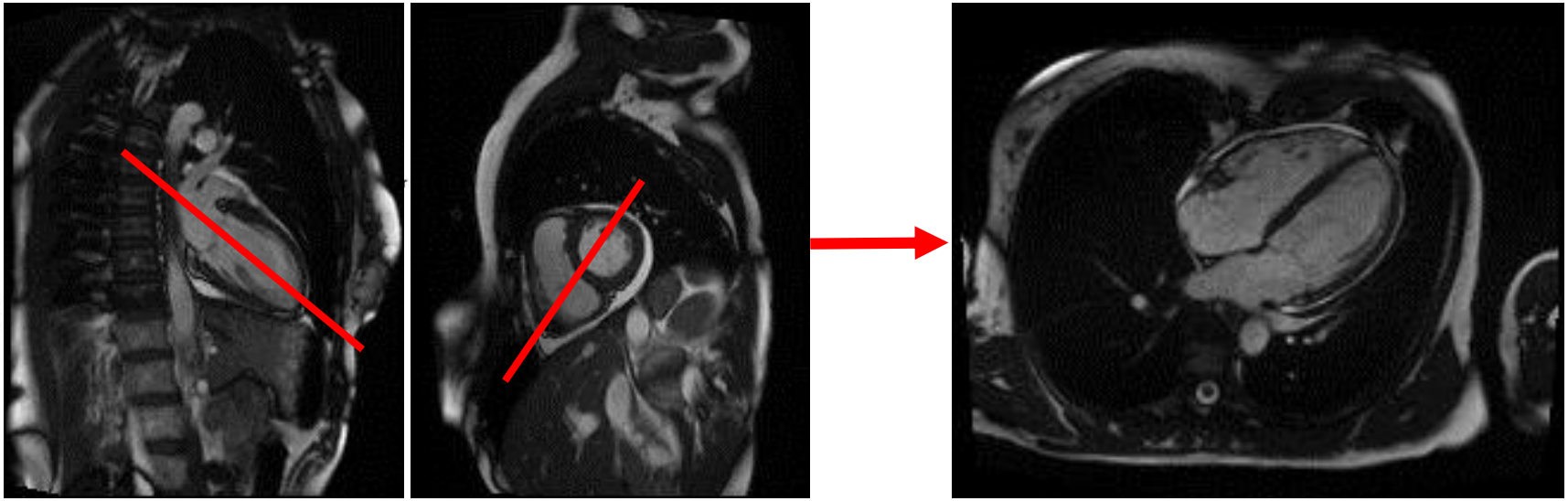


Motion Artefact

* Oksuz et al., ISBI 2018 $Oksuz et al., MICCAI 2018

4-chamber cine cardiac MR

- Good 4-chamber CMR image shows all chambers clearly, enables right and left atrium analysis
- Planned using 2-chamber and short axis images
- Mistakes in planning lead to ‘off axis’ images

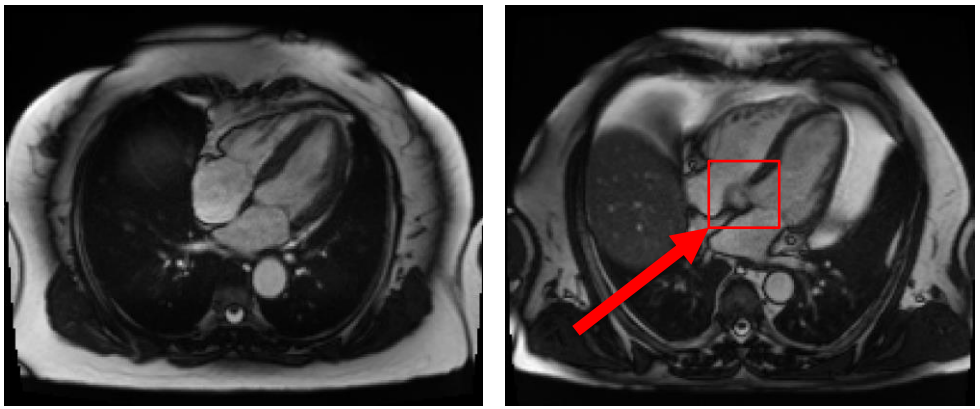


2-chamber view Short axis view 4-chamber view

Oksuz et al., ISBI 2018

Left ventricular outflow tract (LVOT)

- Off-axis acquisitions often show the Left Ventricular Outflow Tract (LVOT)
- Challenges RA and LA analysis
- Automatic LVOT detection can assist automatic quality control/planning



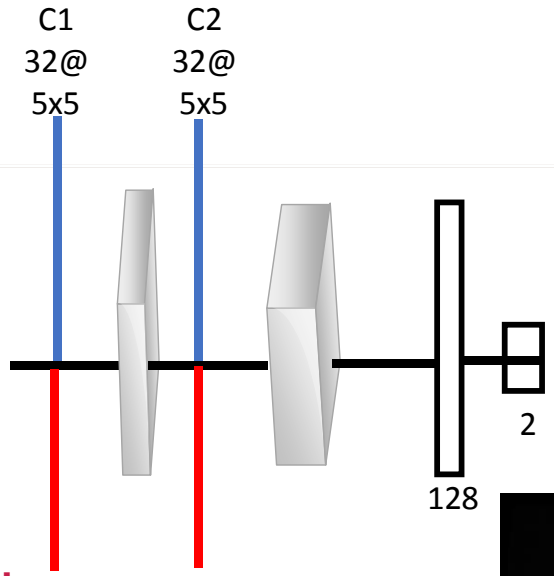
Good Planning

Bad Planning

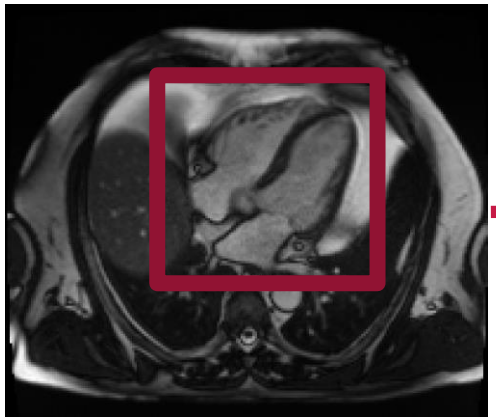
Oksuz et al., ISBI 2018

Method

1. Contrast Normalisation
2. Region of Interest Extraction
3. Training a CNN Model



Input: 2D 4chamber cardiac MR

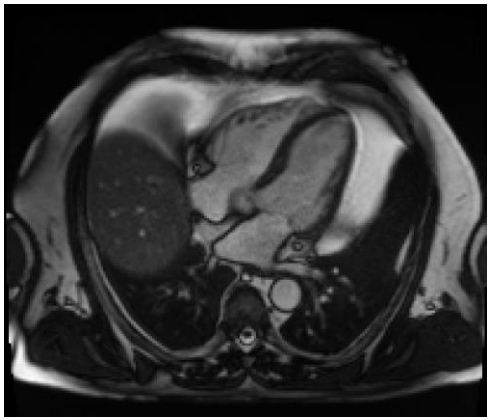


ROI

2D Dataset

CNN Model for Learning

Output: LVOT= 0 or 1



- Similar to Lenet* Model
- Dropout 0.5 after each layer
- ReLU Activation

Oksuz et al., ISBI 2018 *** LeCun et al., Proc IEEE, 1998**

Experimental results

Dataset:

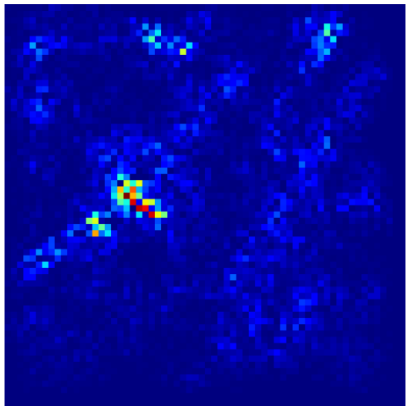
- 123 Good Quality Image and 123 LVOT images from UK Biobank
- 5 temporal frames of each sequence, 615 images for each class

Methods	Accuracy	Precision	Recall
K-Nearest Neighbours	0.613	0.604	0.602
Linear SVM	0.732	0.741	0.736
Decision Tree	0.651	0.626	0.619
Random Forests	0.598	0.613	0.610
Adaboost	0.718	0.729	0.727
Naive Bayesian	0.653	0.625	0.637
Discriminant Analysis	0.669	0.684	0.643
CNN w.o Augmentation	0.801	0.811	0.781
CNN	0.826	0.828	0.821

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

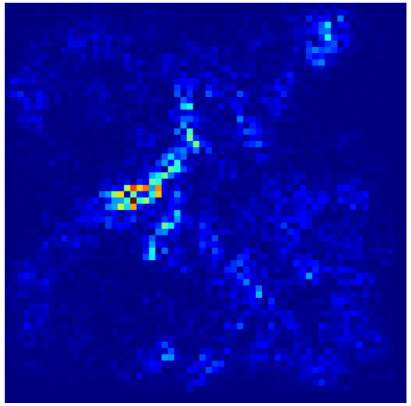
$$\text{Precision} = \frac{TP}{TP + FP} \qquad \text{Recall} = \frac{TP}{TP + FN}$$

LVOT



LVOT attention map*

Good Quality Image



Good Quality Attention Map*

Oksuz et al., ISBI 2018 *** Zhou et al., CVPR, 2016**

Cardiac MR quality issues

1. **Off-axis (4ch)***

Left Ventricular Outflow Tract

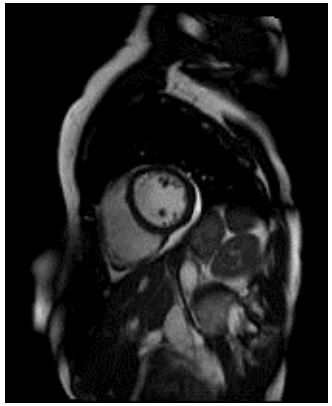
5 chamber look

2. **Motion related issues (SAX)**[$]

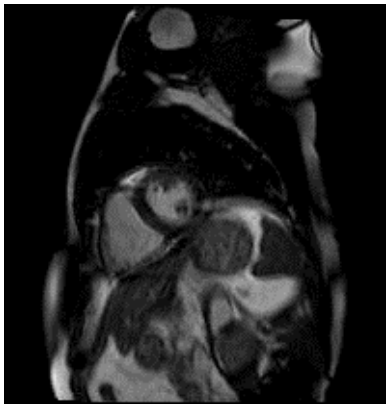
Breathing

Mis-triggering

Arrhythmia



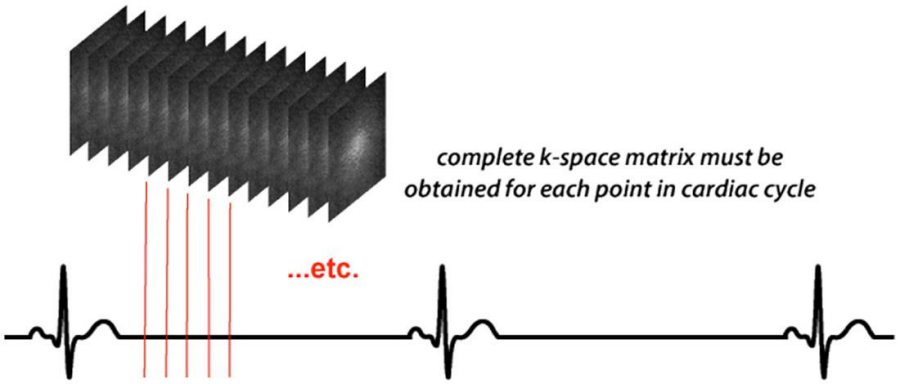
Good Quality

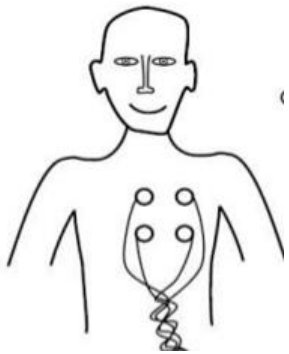


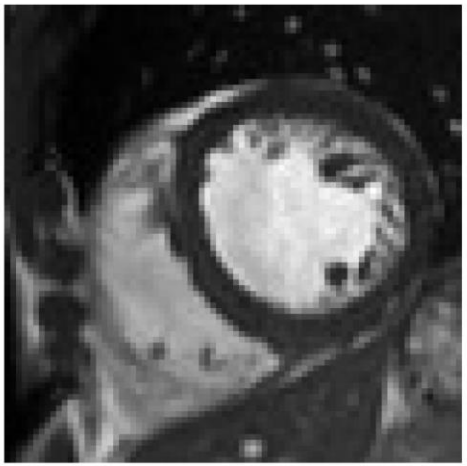
Motion Artefact

[$]Oksuz et al., MICCAI 2018

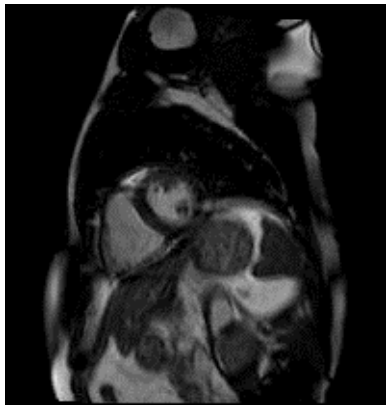
Cardiac cine MR acquisition







Good Quality

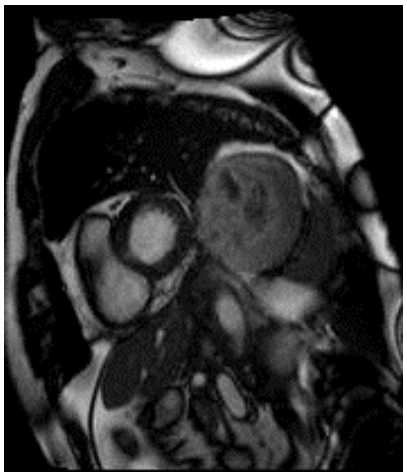


Motion Artefact

Oksuz et al., MICCAI 2018

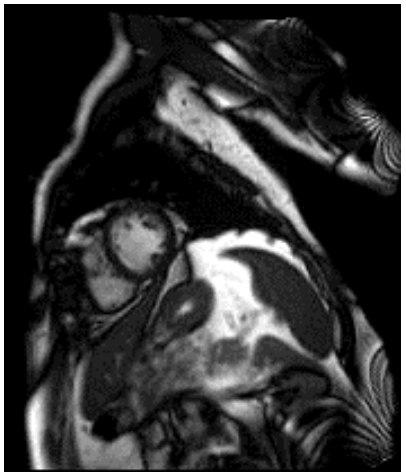
Dataset

- **105 subjects** with motion artefacts (breathing, mistriggering, arrhythmia)
- 53 for mistriggering, 23 for breathing, 24 arrhythmia, 4 mixed
- **105 artefact images, 3360 good quality images**
- **DATA IMBALANCE …**

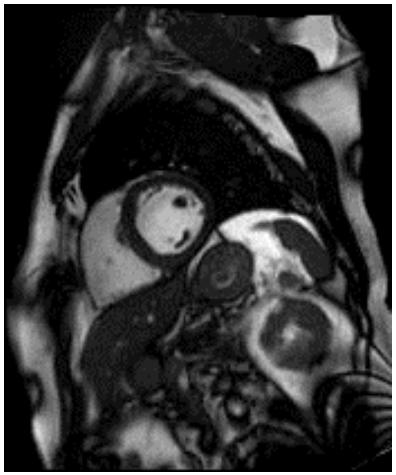


Arrhythmia

Breathing



Mistriggering



Good Quality

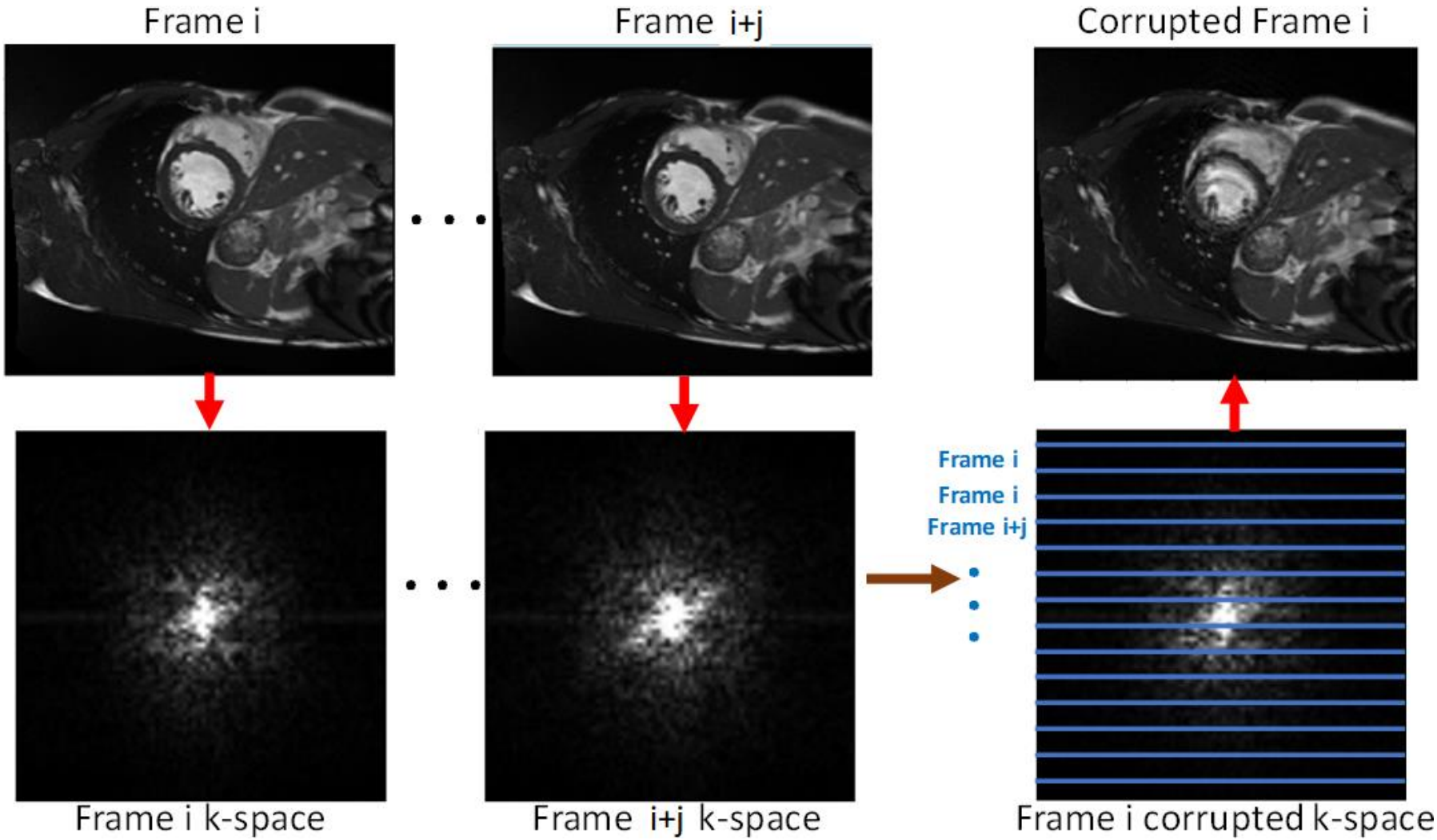
Oksuz et al., MICCAI 2018

Data imbalance

1. **More data**
- Difficult task in many medical imaging applications.
- Not plausible to generate more real low quality medical data.
2. **Resampling dataset**
- Add copies of instances from the under-represented class.
- Delete some data from the over-represented class.
3. **Generate synthetic samples**
- Generate synthetic examples that best represent the original data from the under-represented class.

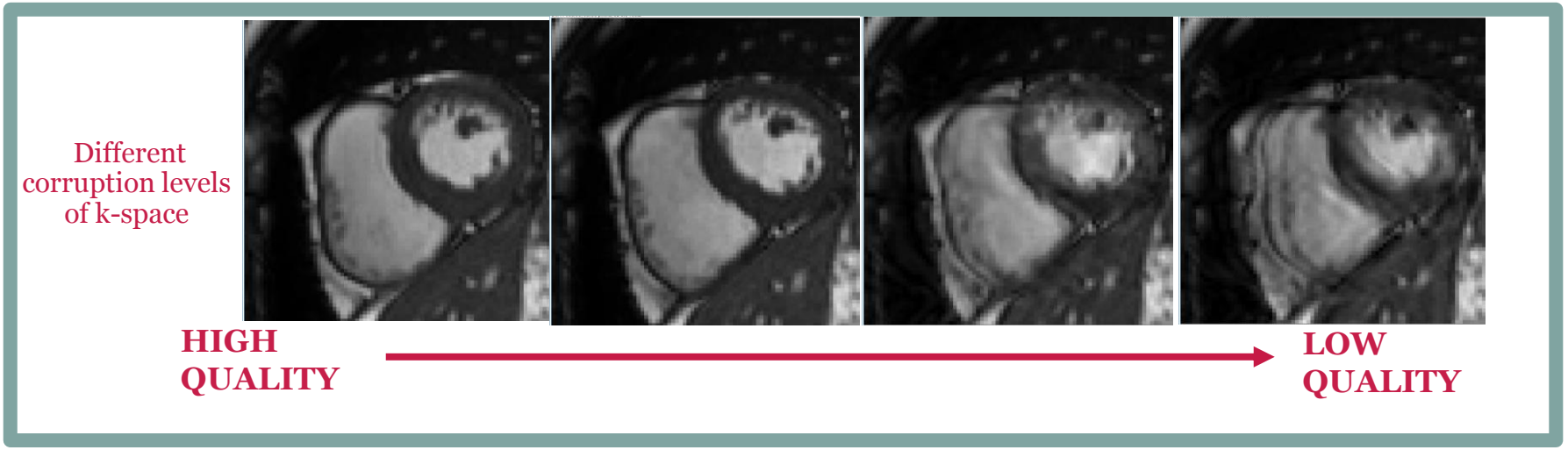
Oksuz et al., MICCAI 2018

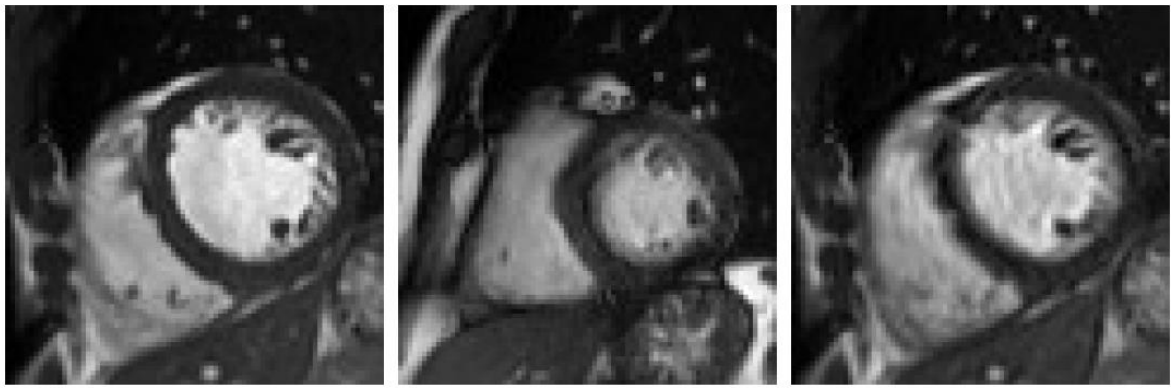
K-space corruption



Oksuz et al., MICCAI 2018

Synthetic images

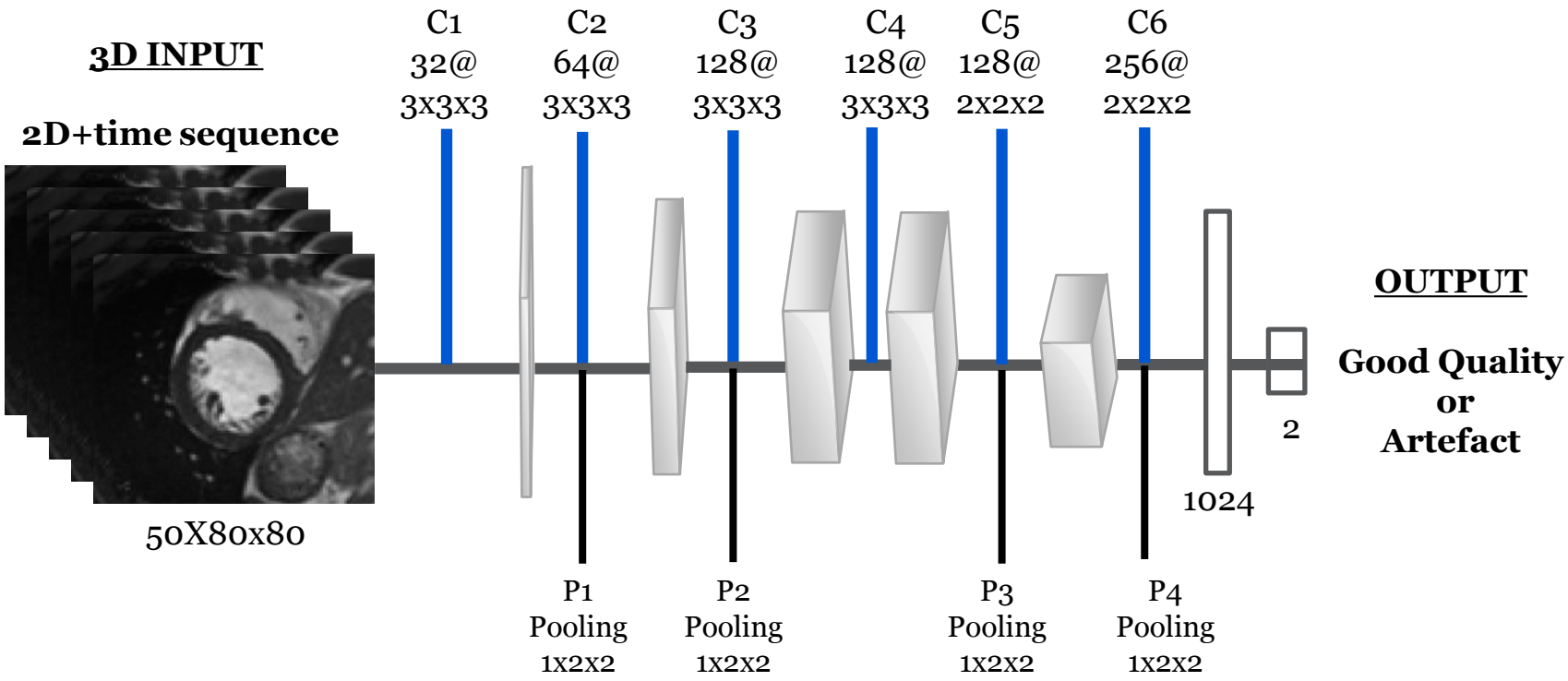




Good quality | Motion artefact | K-space corrupted image

Oksuz et al., MICCAI 2018

3D CNN model



Oksuz et al., MICCAI 2018

Experimental Results

Dataset:

- 105 Artefact Images, 3360 Good quality Images
- Dataset balanced with augmentation

$$\text{Precision} = \frac{TP}{TP + FP} \qquad \text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

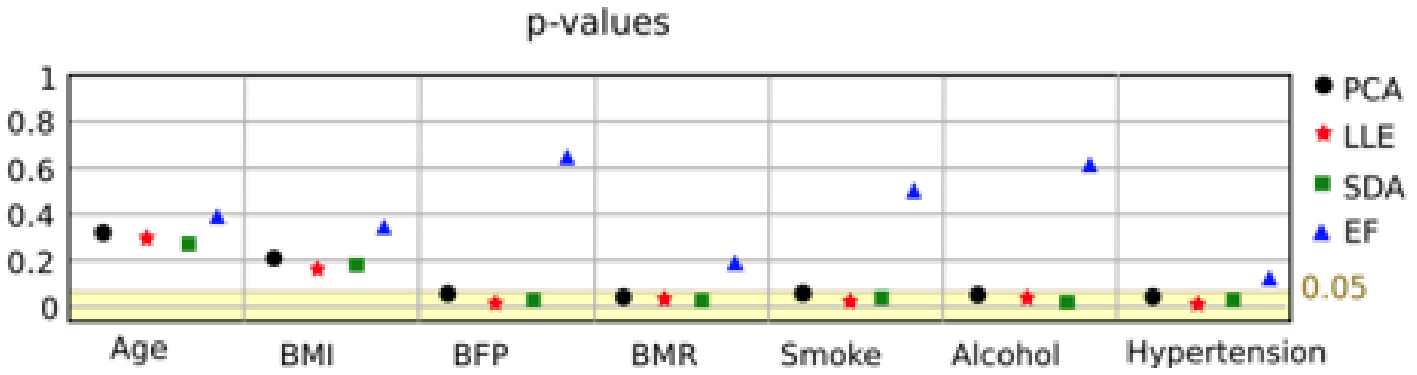
$$\text{F1 score} = \frac{2 * (\text{Recall} * \text{Precision})}{\text{Recall} + \text{Precision}}$$

Methods	Accuracy	Precision	Recall	F1-score
K-Nearest Neighbours	0.952	0.074	0.268	0.116
Linear SVM	0.968	0.721	0.385	0.502
Decision Tree	0.951	0.250	0.385	0.303
Random Forests	0.958	0.320	0.315	0.317
Adaboost	0.960	0.230	0.567	0.327
Naive Bayesian	0.801	0.527	0.183	0.111
Variance of Laplacian	0.958	0.113	0.161	0.133
NIQE *	0.958	0.210	0.248	0.227
CNN with no augmentation	0.968	0.700	0.466	0.560
CNN with translational augmentation	0.974	0.750	0.600	0.667
CNN with k-space augmentation	0.977	0.779	0.642	0.704
CNN with k-space+translational augmentation	0.982	0.809	0.652	0.722

Oksuz et al., MICCAI 2018 *** Mittal et al., IEEE Signal Processing Letters, 2013**

Use of UK biobank data for analysis of factors influencing cardiac health

STACOM poster: Puyol-Anton, et al. “Learning associations between clinical information and motion-based descriptors using a large scale MR-derived cardiac motion atlas”



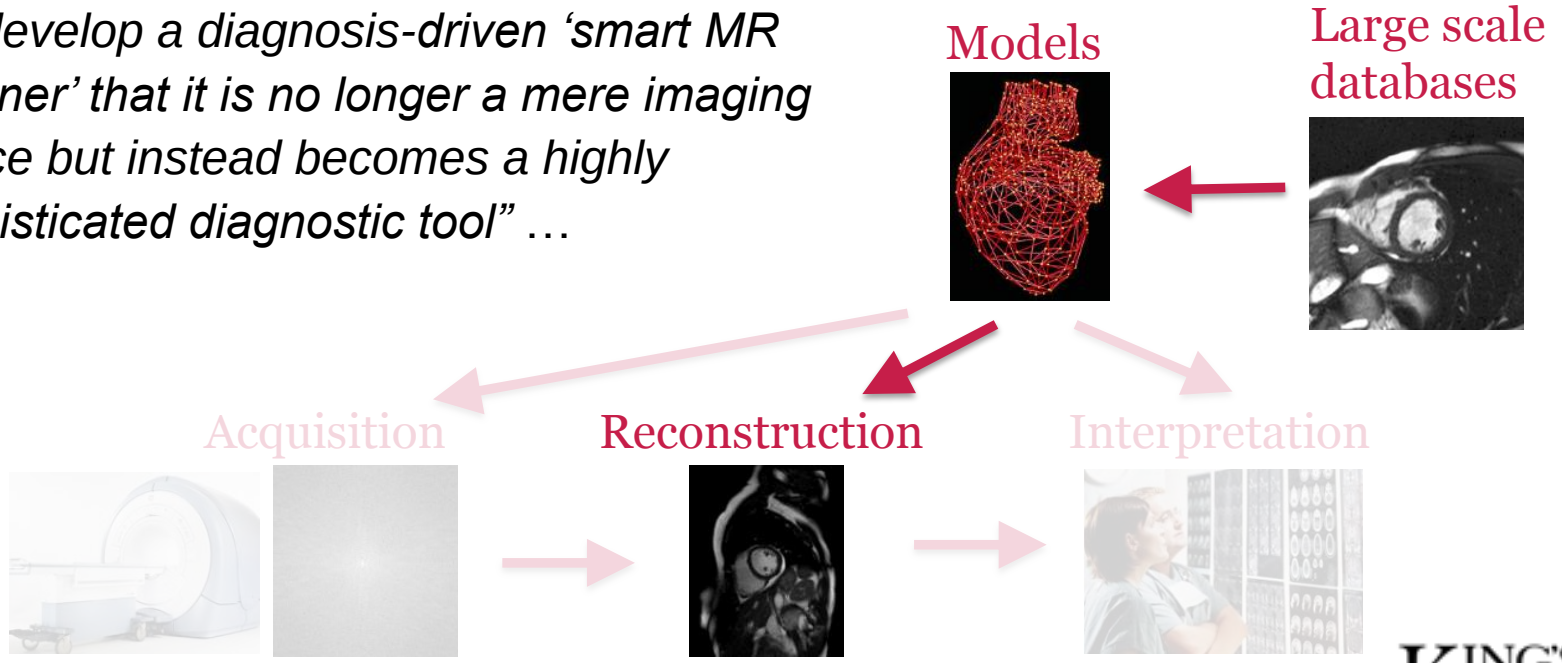
Puyol-Anton et al., MICCAI STACOM 2018

SmartHeart project

5 year grant involving:

- King's College London
- Imperial College London
- Queen Mary University London
- Oxford University

"To develop a diagnosis-driven 'smart MR scanner' that it is no longer a mere imaging device but instead becomes a highly sophisticated diagnostic tool" …



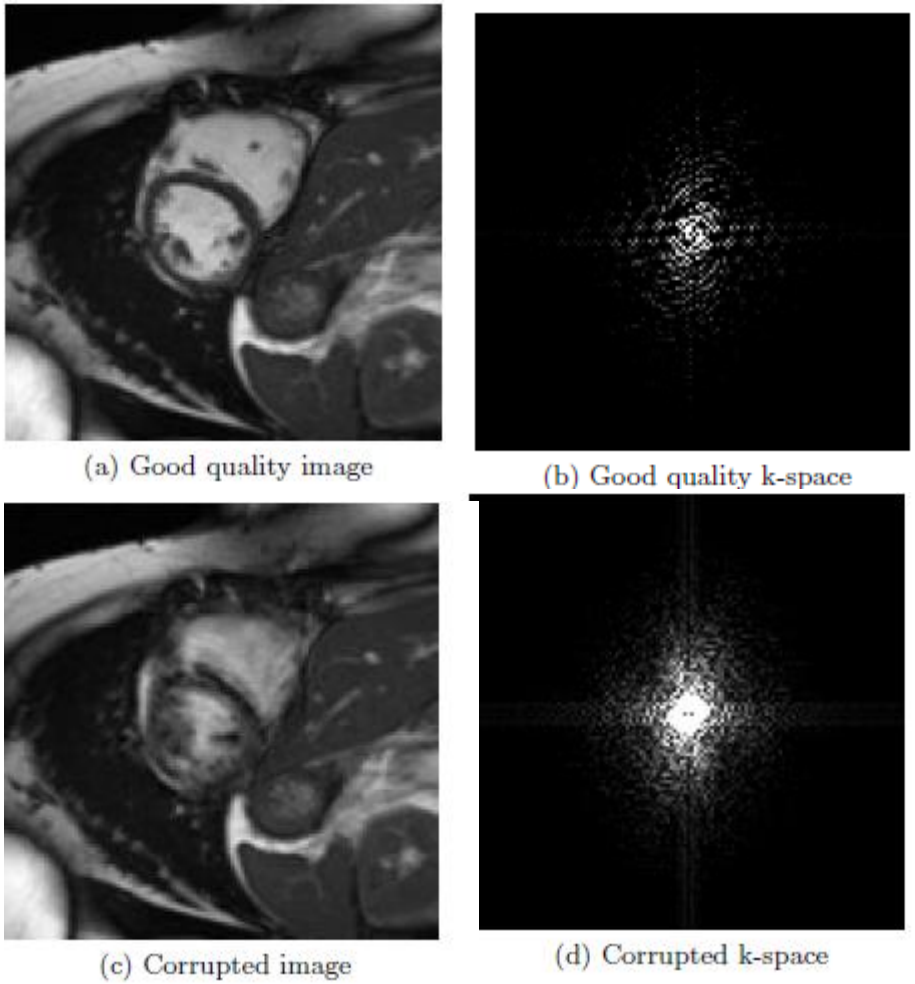
4. Machine learning for robust MR reconstruction

Aim:

MR motion artefact correction during reconstruction

Approaches:

- Denoising in image space (c to a)
- Denoising in k-space (d to b)
- **End-to-end (d to a)***



(a) Good quality image

(b) Good quality k-space

(c) Corrupted image

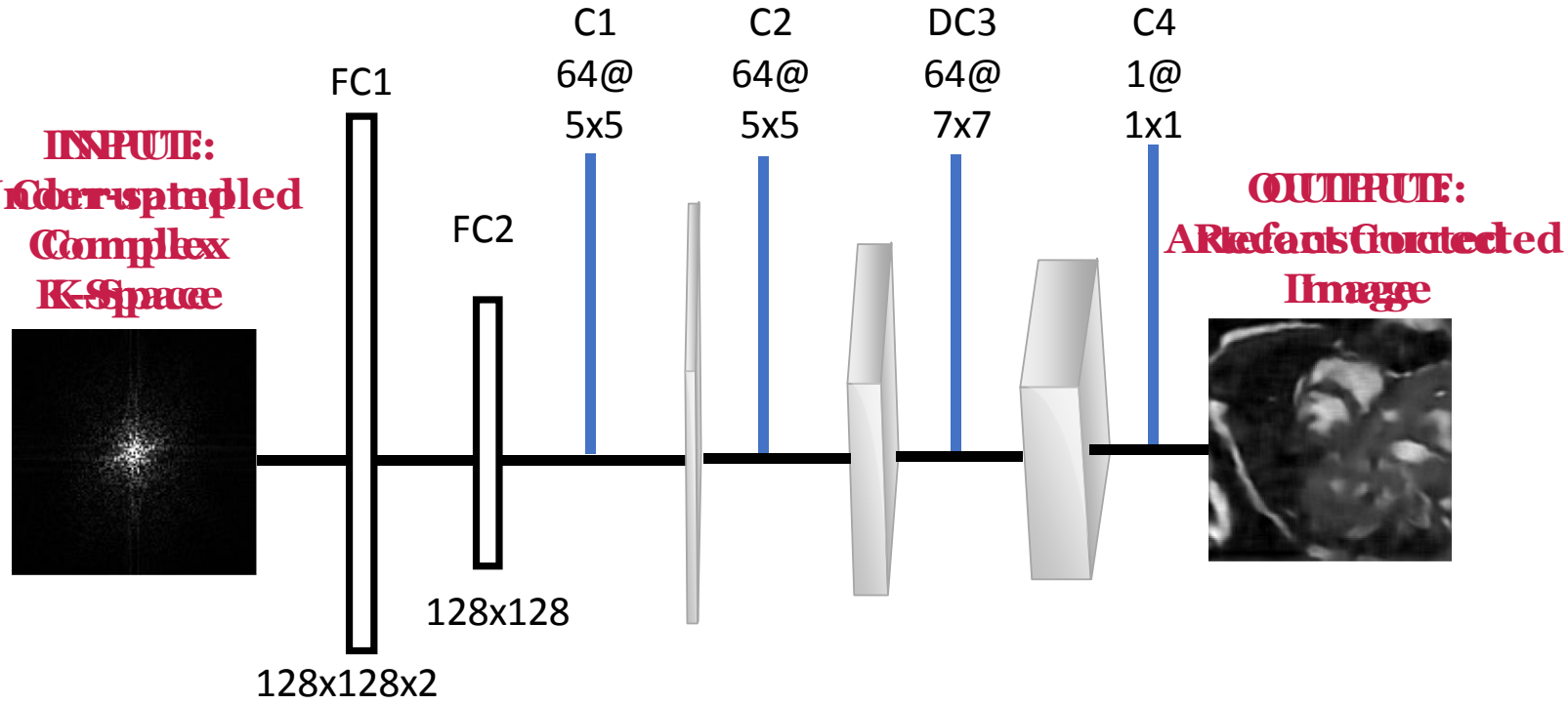
(d) Corrupted k-space

Oksuz et al., MICCAI MLMIR, 2018

Automap

Image reconstruction by domain transform manifold learning

- Developed for high quality image reconstruction from under-sampled k-space
- Insufficient image quality for corrupted k-space



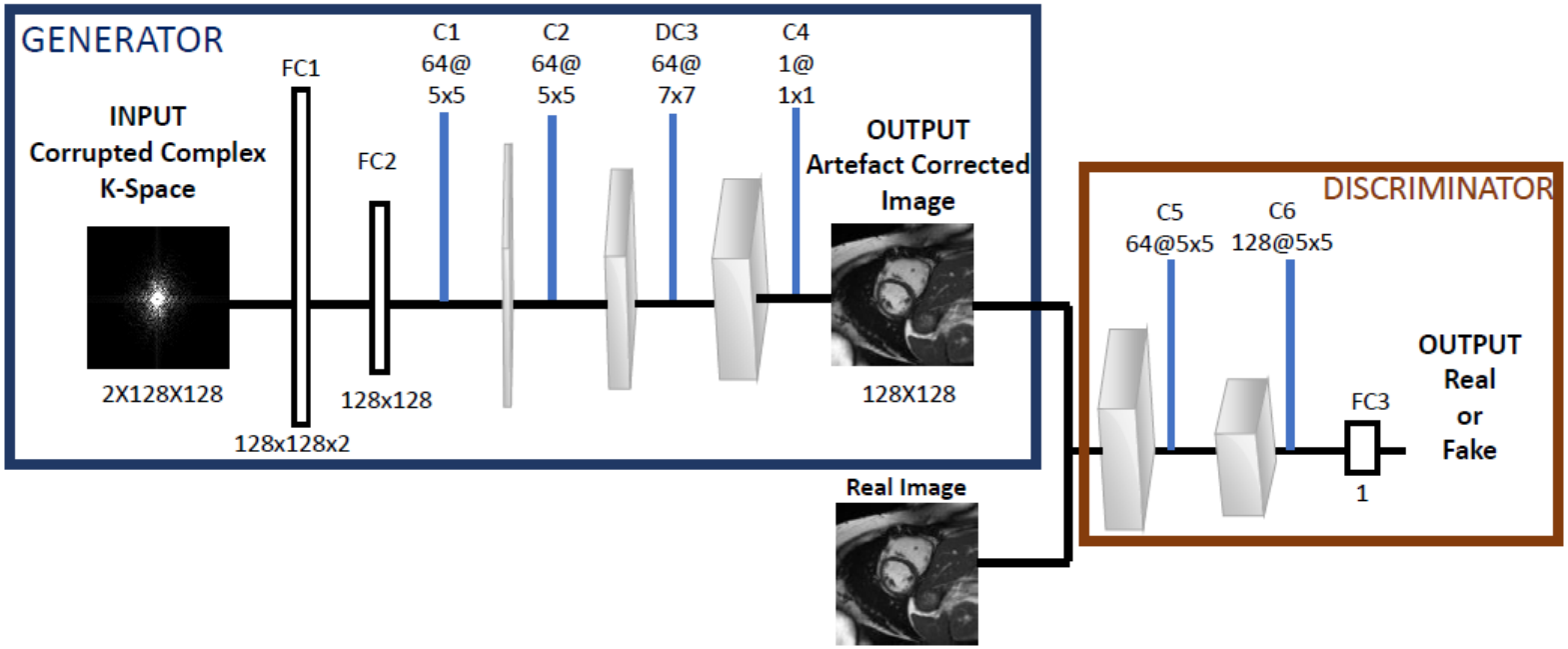
Oksuz et al., MICCAI MLMIR, 2018

* Zhu et al., Nature, 2018

Automap-GAN setup

Adversarial setup for motion artefact correction

- Improved robustness and deblurring of the image outputs



Oksuz et al., MICCAI MLMIR, 2018

Experimental results

Dataset:

- Synthetically generated corruptions
- 75000 2D images for training, 2500 for testing

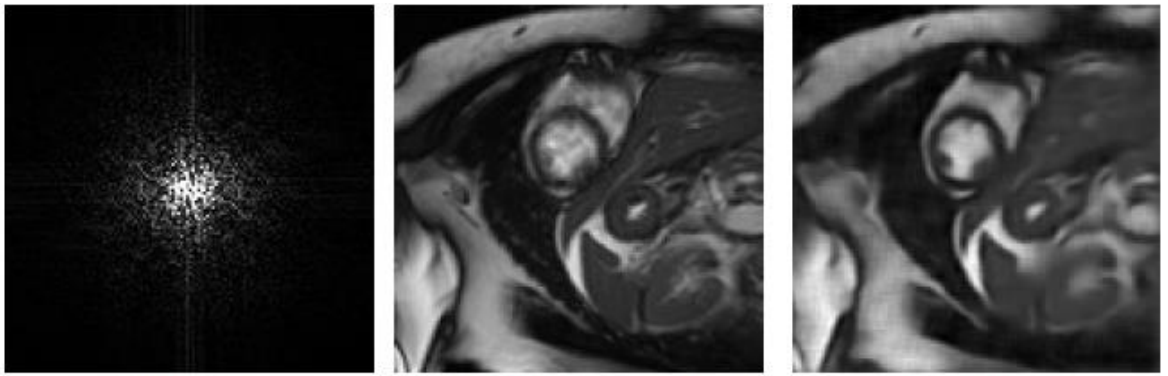
Methods	RMSE	PSNR	SSIM
Inverse Fourier Transform	0.045	27.8	**0.883**
Proposed-ImageNET	0.032	31.1	0.766
Automap-Cardiac	0.029	32.7	0.814
Proposed-Cardiac	**0.027**	**35.1**	0.850

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_x N_y}\sum_{x=0}^{N_x}\sum_{y=0}^{N_y}(r(x,y)-p(x,y))^2}$$

$$\mathrm{PSNR} = 20log_{10}\left(\frac{\sum_{x=0}^{N_x}\sum_{y=0}^{N_y} r(x,y)^2}{\sqrt{\sum_{x=0}^{N_x}\sum_{y=0}^{N_y}(r(x,y)-p(x,y))^2}}\right)$$

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

In-vivo example



(a) K-space (b) Motion corrupted image (c) Proposed

Oksuz et al., MICCAI MLMIR, 2018

QUESTIONS ??

Ilkay Oksuz

Esther Puyol Anton

Bram Ruijsink

Devis Peressutti

Matt Sinclair

Claudia Prieto

Julia Schnabel

Rene Botnar

andrew.king@kcl.ac.uk

@atoandyking

Group web site: https://kclmmag.org/
SmartHeart project: https://wp.doc.ic.ac.uk/smartheart/
UK Biobank: http://www.ukbiobank.ac.uk/

biobank uk
Improving the health of future generations

(Application 17806)

smartheart

EPSRC
Pioneering research and skills

KING'S College LONDON