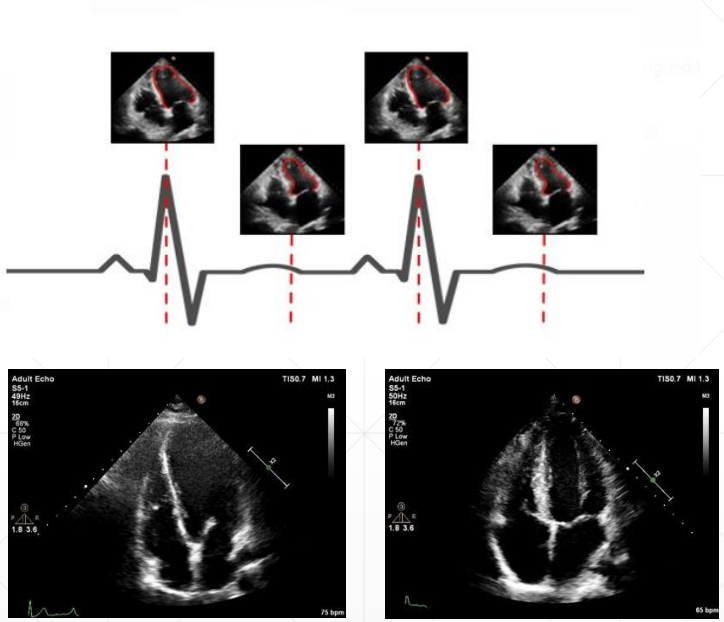
Echocardiographic Phase Detection Using Neural Networks

Beth Lane, PhD Candidate
School of Computing and Engineering, University of West London

Accurate identification of end-diastolic (ED) and end-systolic (ES) frames in echocardiographic cine loops is essential when measuring cardiac function. Manual selection by human experts is challenging and error prone.

We present a deep neural network trained and tested on multi-centre patient data for accurate phase detection in apical four-chamber videos of arbitrary length, spanning several heartbeats, with performance indistinguishable from that of human experts.



MIDL Lübeck 2021

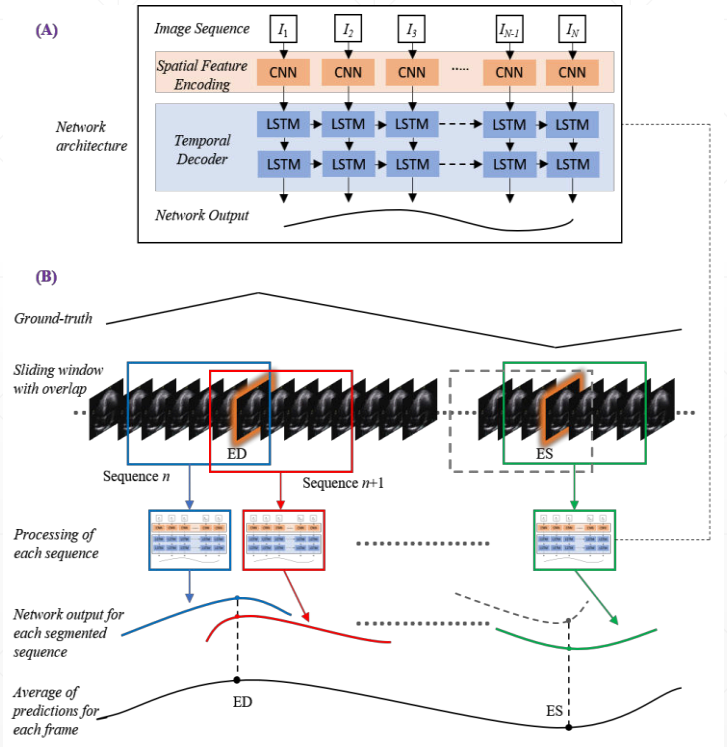
UNIVERSITY OF WEST LONDON
The Career University

intelligentsensing&vision

Three multi-centre apical 4-chamber (A4C) **datasets** were used in this study:

Dataset Name	PACS-dataset	MultiBeat-dataset	EchoNet-dataset
Source	Private NHS Trust PACS Archives - Imperial College Healthcare	Private St Mary's Hospital Acquired for this study	Publicly available Stanford University Hospital echonet.github.io/dynamic
Number of videos/patients	1,000	40	10,030
Length of videos	1-3 heartbeats	≥ 10 heartbeats	1 heartbeat
Ground-truth	2 annotations by 2 experts	6 annotations by 5 experts (twice by one expert)	1 annotation
Original size (pixels)	(300-768)×(400-1024)	422×636	112×112
Frame rate (fps)	23-102	52-80	50
Format	DICOM	DICOM	AVI
Use	**Training/Testing**	**Testing**	**Testing**

The **model architecture** comprises a CNN unit (ResNet50 with ImageNet weights) for encoding spatial information, a RNN (2x LSTM) unit for decoding temporal information, from which return sequence is set to true then flattened and regressed through a Dense layer in chunks of 30 frames. A fixed-stride of 1 frames sliding window allows multiple predictions to be averaged for each input frame. A peak finding algorithm identifies discrete predictions for ED and ES relative to a predetermined threshold.



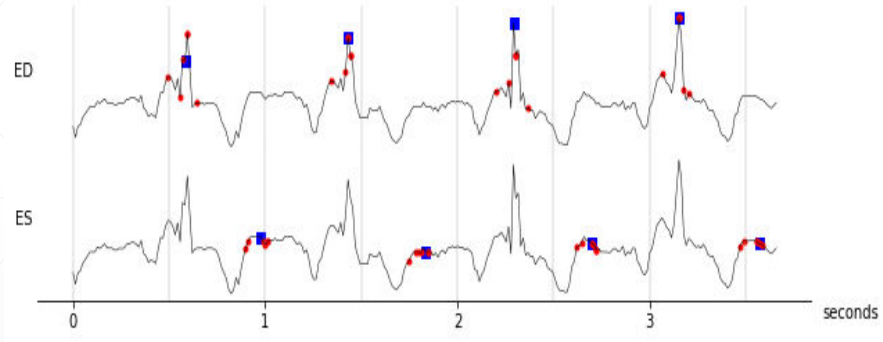
Average absolute frame difference (aaFD) notation was used as the **evaluation metric**:

$$aaFD = \frac{1}{N}\sum_{t=1}^{N}|y^t - \hat{y}^t|,$$

PACS-dataset results:

Model/Operator	ED		ES		Detection Time (s)
	aaFD	μ ± σ	aaFD	μ ± σ	
ResNet50 + 2x-LSTM	0.66	-0.09±1.10	0.81	0.11±1.29	0.776±0.33
Operator-2 (inter-observer)	1.55	-1.35±1.31	1.44	-0.90±1.80	26±11

The figure below represents an extracted ECG trace from the Multibeat-dataset with expert annotations as red dots and model predictions as blue squares



Multibeat-dataset results:

Model/Operator	ED		ES	
	aaFD	μ ± σ	aaFD	μ ± σ
Operator-1a vs Operator-1b	1.96	-0.22 ± 2.76	1.90	0.25 ± 3.75
Operator-1a vs Operator-2	2.65	-1.22 ± 4.26	3.67	-2.25 ± 4.68
Operator-1a vs Operator-3	5.82	-5.51 ± 3.77	4.80	-4.46 ± 3.77
Operator-1a vs Operator-4	1.72	-0.87 ± 2.29	2.01	-0.97 ± 3.48
Operator-1a vs Operator-5	3.27	-2.96 ± 2.57	4.11	-3.64 ± 3.67
Operator-1a vs model	2.62	-1.34 ± 3.27	1.86	-0.31 ± 3.37

EchoNet-dataset results:

Model/Operator	ED		ES	
	aaFD	μ ± σ	aaFD	μ ± σ
Model	2.30	0.16±3.56	3.49	2.64±3.59