Convolutional Neural Net Transformer (CNNT) for free-breathing real-time cine imaging

A novel architecture for real-time cine imaging at higher accelerations

Azaan Rehman, Rhodri H. Davies, Iain Pierce,
Marianna Fontana, James C. Moon, Peter Kellman, Hui Xue

6th May 2022

SCMR
Society for Cardiovascular Magnetic Resonance

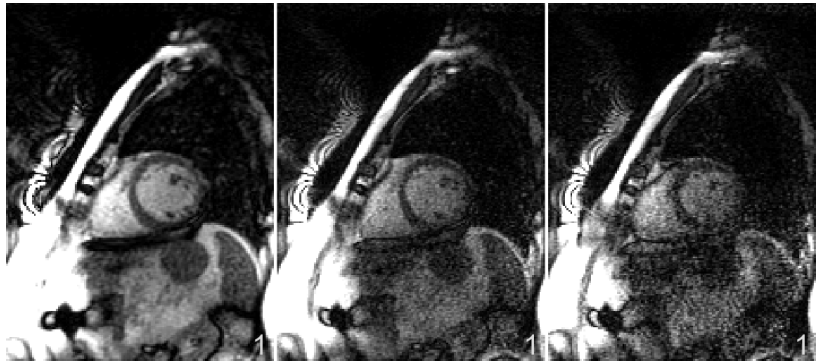
EACVI
European Association of Cardiovascular Imaging
European Society of Cardiology

Declaration of Interests

- **No conflicts to declare**

Background: Task

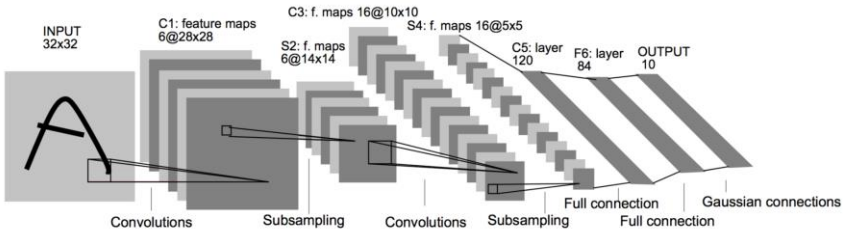
- Real-time cine imaging
  - Does not require breath-holding
  - Can be used in the presence of irregular heartbeats
- To maintain good temporal and spatial resolution, higher acceleration (4x,5x,6x parallel imaging) is required

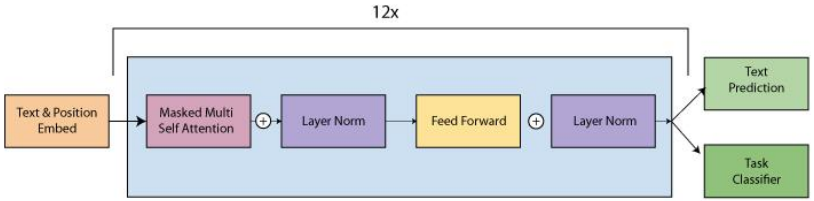


Background: CNN + Transformers

- **Convolution Neural Nets**
  - Commonly used for 2D data (images)
  - Efficient with parameters
  - Good in reading spatial dependencies
  - Limited learning of temporal relations

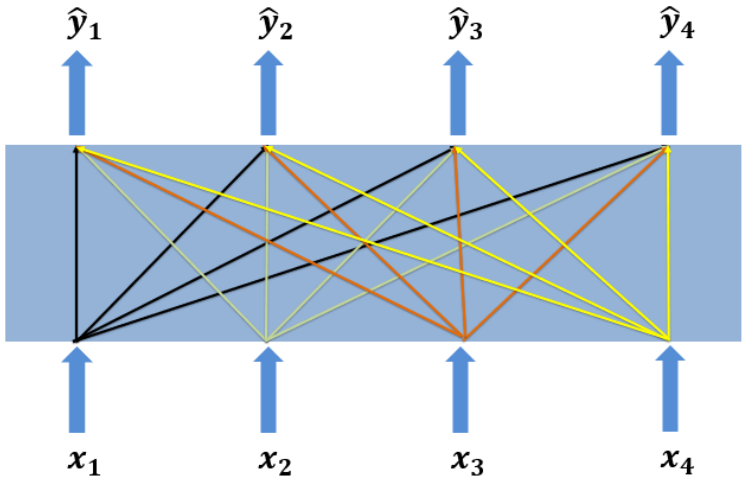
- **Transformers**
  - Commonly used with sequence data
  - Can capture long range temporal dependencies
  - Uses attention mechanism for faster training across time





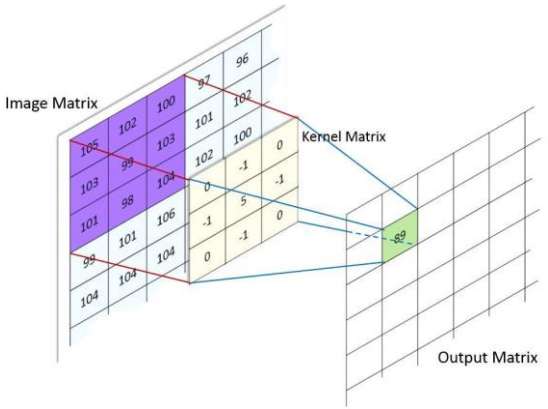
Attention in transformers

- **Baseline and most SotA transformers use linear layers**
  - To map inputs to learnable matrices (Key, Query and Value)



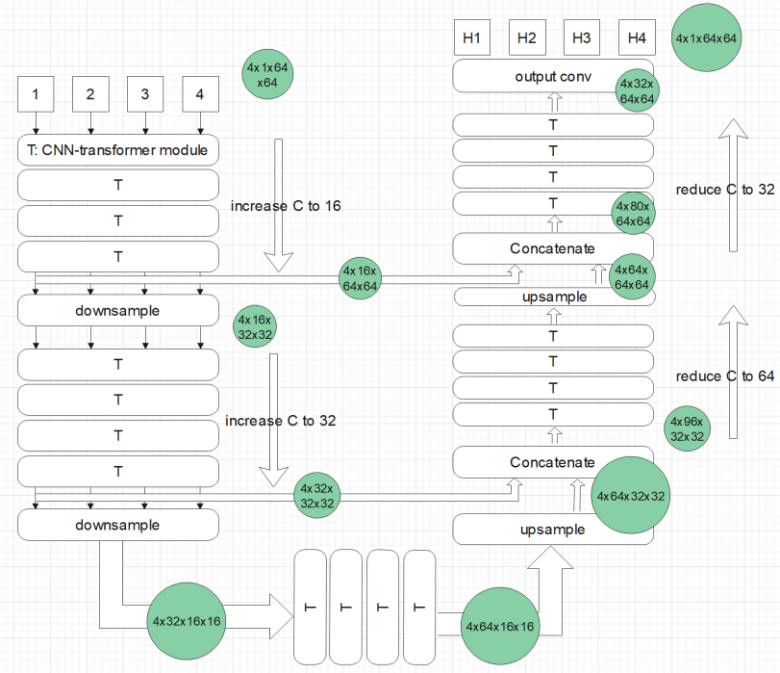
Linear to Conv attention

- **We propose: replace linear layer mappings to convolution layers instead**
  - Heavily reduces the effective parameters
  - Makes use of learning spatial correlations as CNNs
  - Can control the number of channels
  - CNNT layer (Convolutional Neural Network Transformer layer)

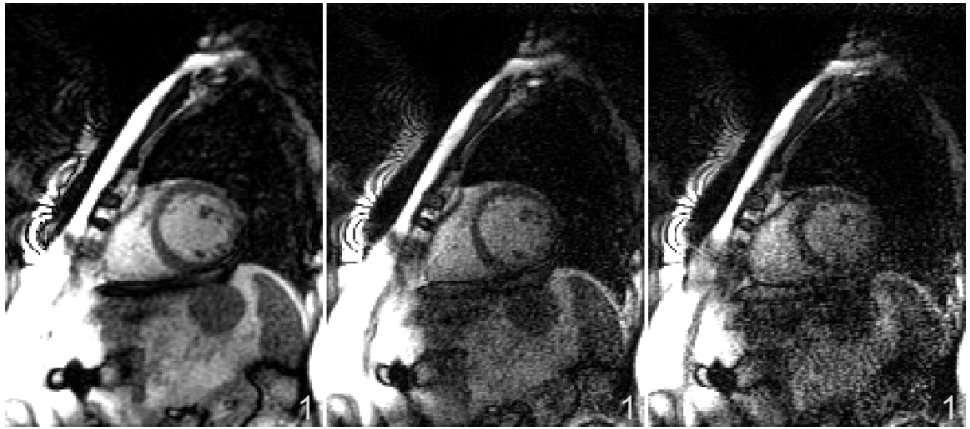


CNNT Unet

- **Unets have proven to work well in medical imaging**
  - Use up/down sampling to control spatial resolution
  - Skip connections for long-term gradient flow
- **Stack CNNTs per block**
  - Takes 2D+T
  - Outputs 2D+T
  - Variable sized inputs



Data acquisition

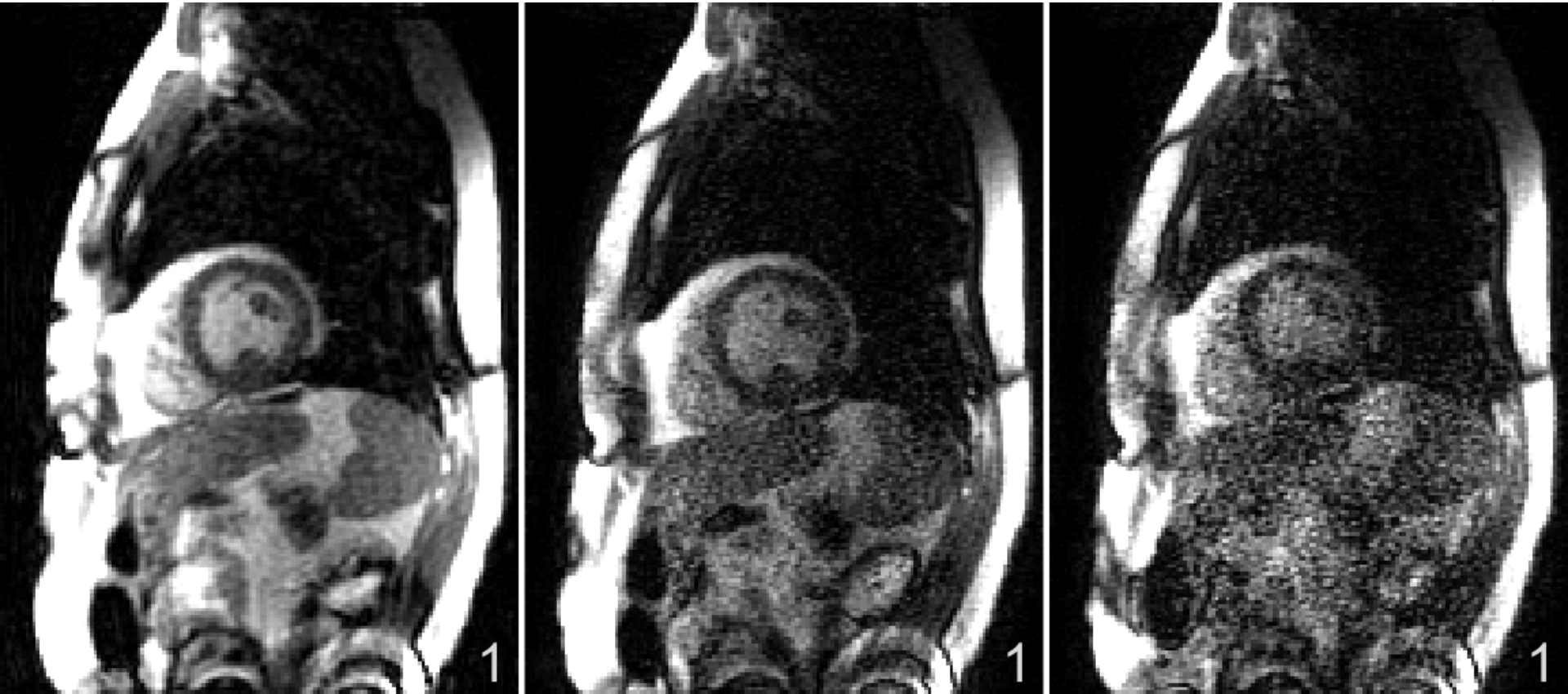


- FOV 360×270mm2, flip angle 50°
- R4 acceleration
  - Matrix size: 160x90
  - Temporal resolution: 40ms
- R5 acceleration
  - Matrix size: 192x108
  - Temporal resolution: 42ms
- R6 acceleration
  - Matrix size: 192x108
  - Temporal resolution: 35ms

Raw data

SCMR
Society for Cardiovascular Magnetic Resonance

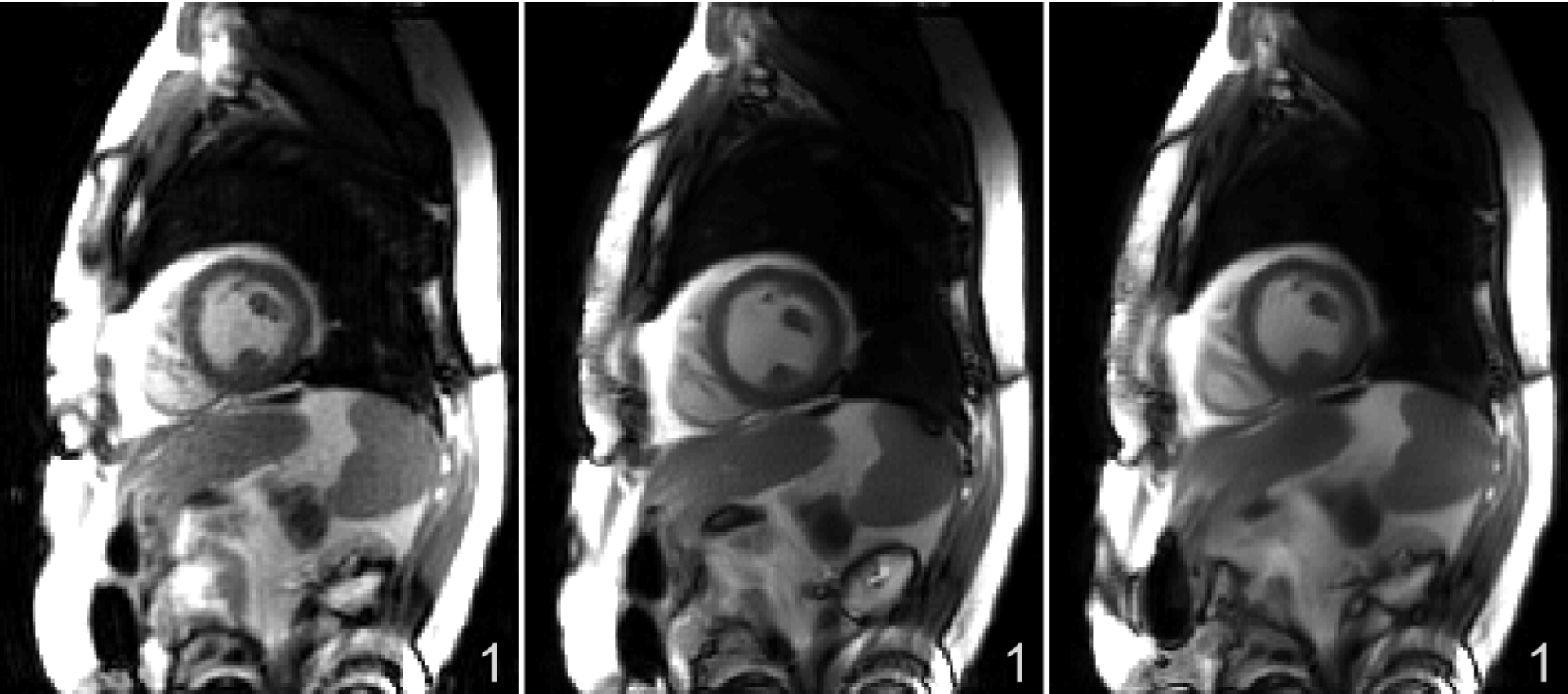
EACVI
European Association of Cardiovascular Imaging



Result: CNNT Unet

SCMR
Society for Cardiovascular Magnetic Resonance

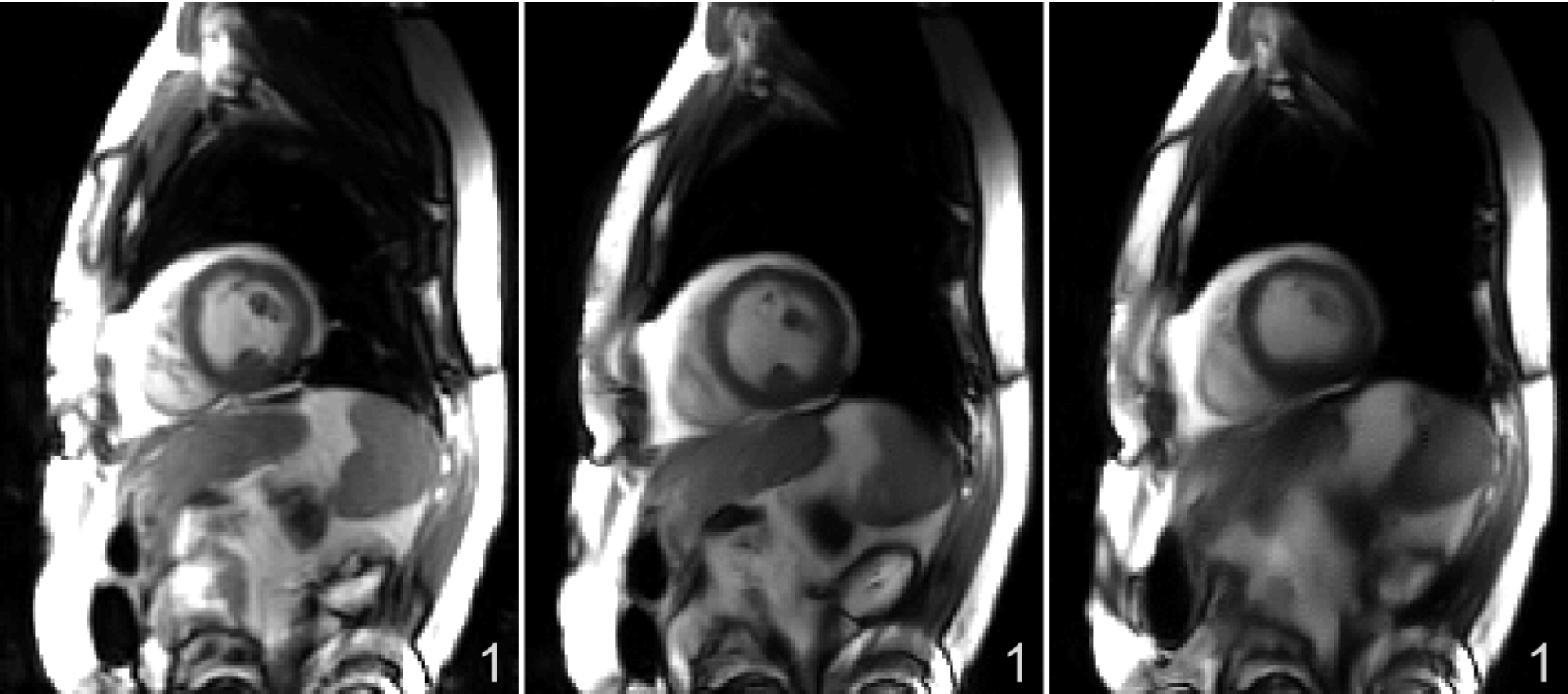
EACVI
European Association of Cardiovascular Imaging



Result: 2D Unet

SCMR
Society for Cardiovascular Magnetic Resonance

EACVI
European Association of Cardiovascular Imaging

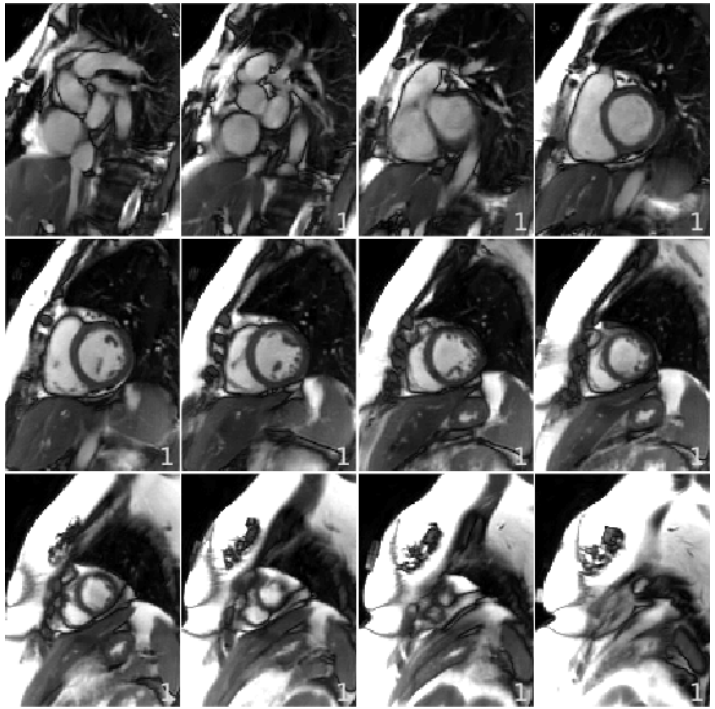


Results

- **Comparison between standard TGRAPPA and CNNT for N=10 patients on Regions of Interest of Blood and Myocardium**

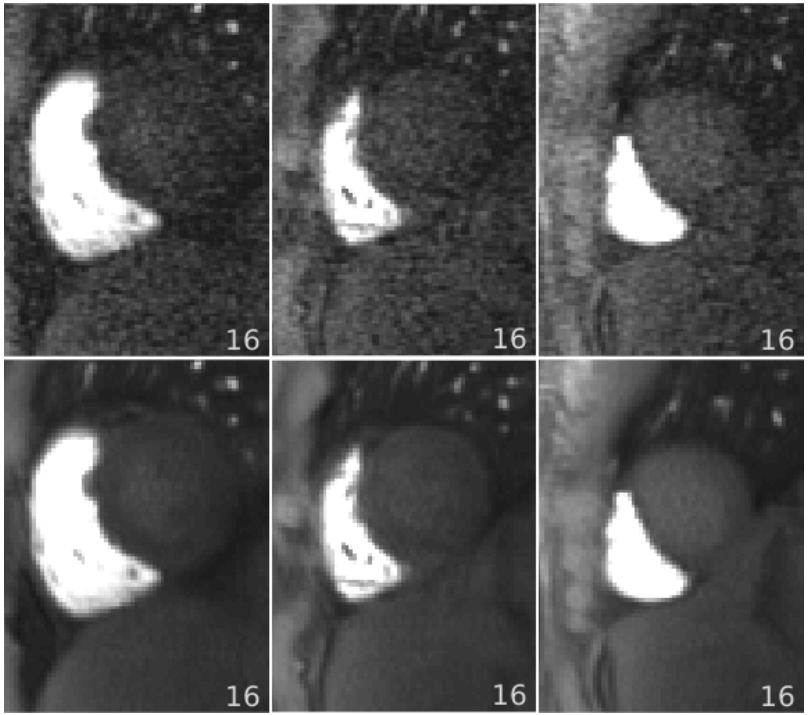
Signal-to-noise ratio	R4		R5		R6	
	TGRAPPA	CNNT	TGRAPPA	CNNT	TGRAPPA	CNNT
Blood	32.4±10.7	54.9±16.0	13.1±4.97	43.9±12.7	8.60±3.08	31.9±8.84
% increase from TGRAPPA	-	**170 %**	-	**335%**	-	**371%**
Myocardium	13.4±4.77	44.9±11.7	5.41±1.91	34.3±8.43	3.49±1.21	28.9±6.44
% increase from TGRAPPA	-	**335 %**	-	**634%**	-	**828%**
p-value	<1e-5 for both blood and myocardium		<1e-5		<1e-5	

Results: RetroCine



- Retro-gated BSSPF Cine
- 30ms temporal resolution
- R=4 acceleration
- Matrix size 256x144, FOV 360x270mm2
- 3s breath-hold per slice

Results: High Resolution Perfusion



- Free-breathing perfusion
- 1.5T scanner
- Spatial resolution 256x144
- FOV 360x270m2
- R=4 acceleration

Conclusion

- **We propose a novel architect: Convolution Neural Network Transformer (CNNT)**
- **CNNT Unet outperforms current practices by a significant margin in the image enhancement task**
- **Suitable for temporal CMR imaging**
- **Opens doors for improvements in other imaging applications**
  - RetroCine, Perfusion, T1/T2 mapping, Flow imaging etc.