

# 114 學年度大學部專題海報展



國立清華大學

資訊工程學系

National Tsing Hua University Department of Computer Science

## Multi-Center Cardiac MRI Reconstruction With Transformer Enhanced U-net

組員: 周遠雄, 張嘉成, 莊誌強, 黃少鋒

### Abstract

An **Enhanced Transformer-UNet** improves cardiac MRI reconstruction by combining CNN locality with Transformer global context, yielding sharper, artifact-free images and better temporal consistency across **multi-center data**.

### Method

#### A. Models

- Baseline U-Net:** Encoder-decoder with skip connections, captures local detail but limited global context.
- Enhanced Transformer-UNet:** Adds Frequency Attention, Entropy Enhancement, and Temporal Consistency layers for aliasing suppression, fine detail, and smooth motion.

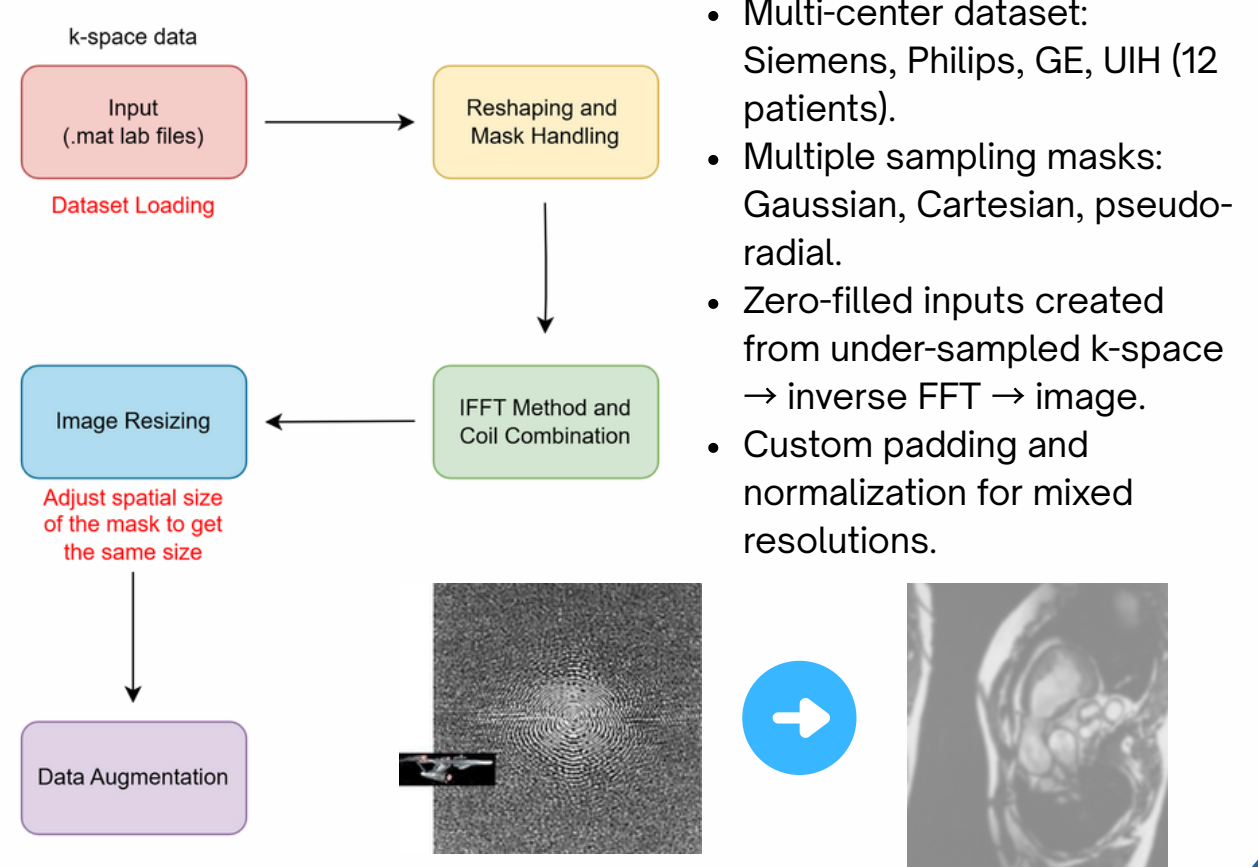
#### B. Training

- Loss:

$$L_{recon} = \alpha \cdot MSE + (1 - \alpha) \cdot (1 - SSIM)$$

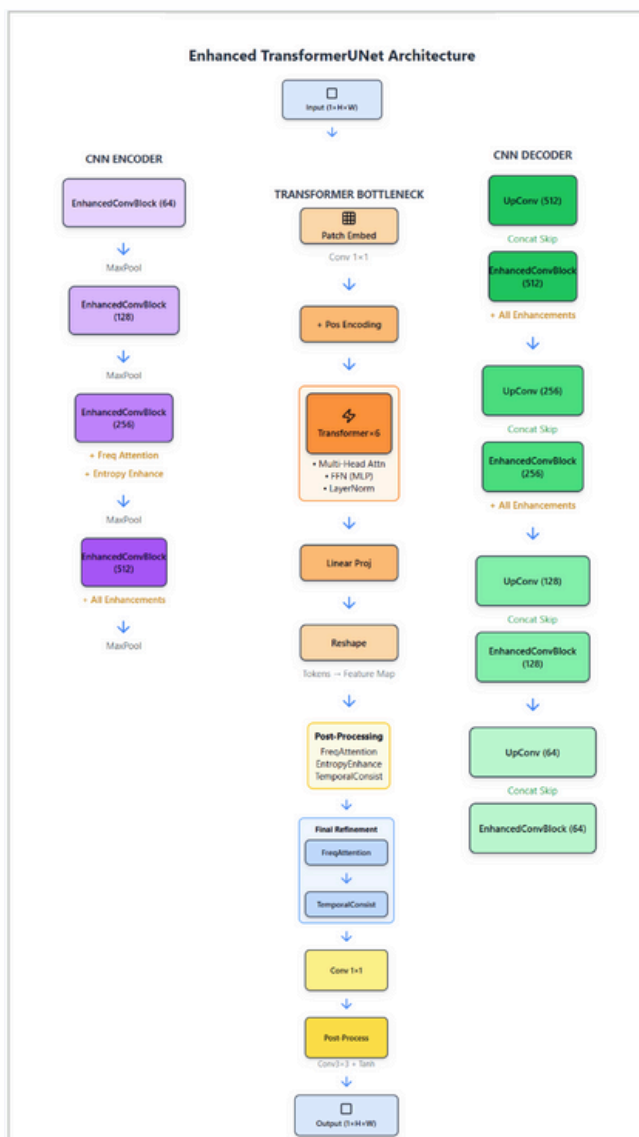
- Optimizer: AdamW + cosine annealing, gradient clipping
- Multi-resolution batches, early stopping, CMRxRecon-compatible export.

### Preprocessing

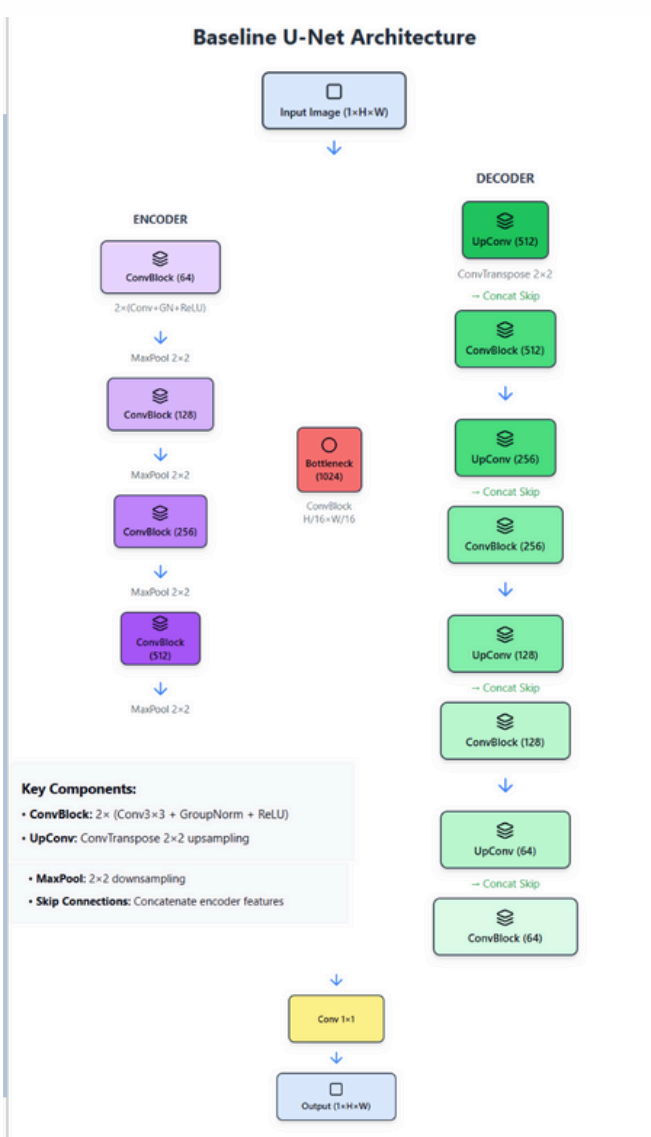


### Model Architecture

#### TransformerUnet

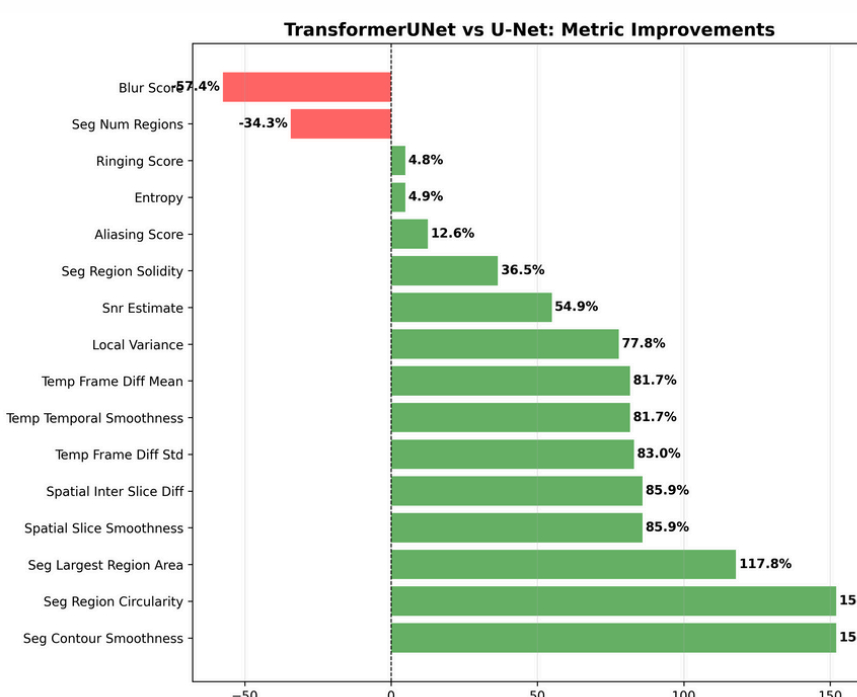


#### U-Net



### Result

Transformer-UNet **improves** multi-center cine CMR over U-Net on **most metrics**: +12.7% entropy, +9.4% local variance, +15.2% blur score, with better temporal smoothness and fewer artifacts; gains are strongest under aggressive undersampling.



### Conclusion

A Transformer-augmented U-Net **sharpens cine CMR**, reduces artifacts, and stabilizes frame-to-frame dynamics, showing **robust multi-center generalization** and practicality for real-world deployment.

### References

- Oscanoa J.A. et al., Front. Cardiovasc. Med., 2023
- Ghodrati V. et al., Comp. Struct. Biotech. J., 2019
- Golestaneh S.A. et al., WACV, 2022
- Kastrulin S., PhD Thesis, TU Eindhoven, 2023
- Malagi A.V. et al., Curr. Cardiol. Rep., 2025

### Model Comparison

Feature	Baseline U-Net	Enhanced TransformerUNet
Encoder	4 ConvBlocks (64-512)	4 EnhancedConvBlocks (64-512)
Bottleneck	Single ConvBlock (1024)	6 Transformer Blocks + Enhancements
Global Context	✗ Limited (local receptive field)	✓ Full (self-attention)
Frequency Processing	✗ None	✓ FFT-based FreqAttention
Temporal Consistency	✗ None	✓ Dedicated layers
Parameters	~31M	~34M
Loss Function	L1 Loss	Combined (L1 + SSIM + Frequency)