

Deep Learning for Dynamic MRI Reconstruction

BME1312 Artificial Intelligence in Biomedical Imaging

ShanghaiTech University

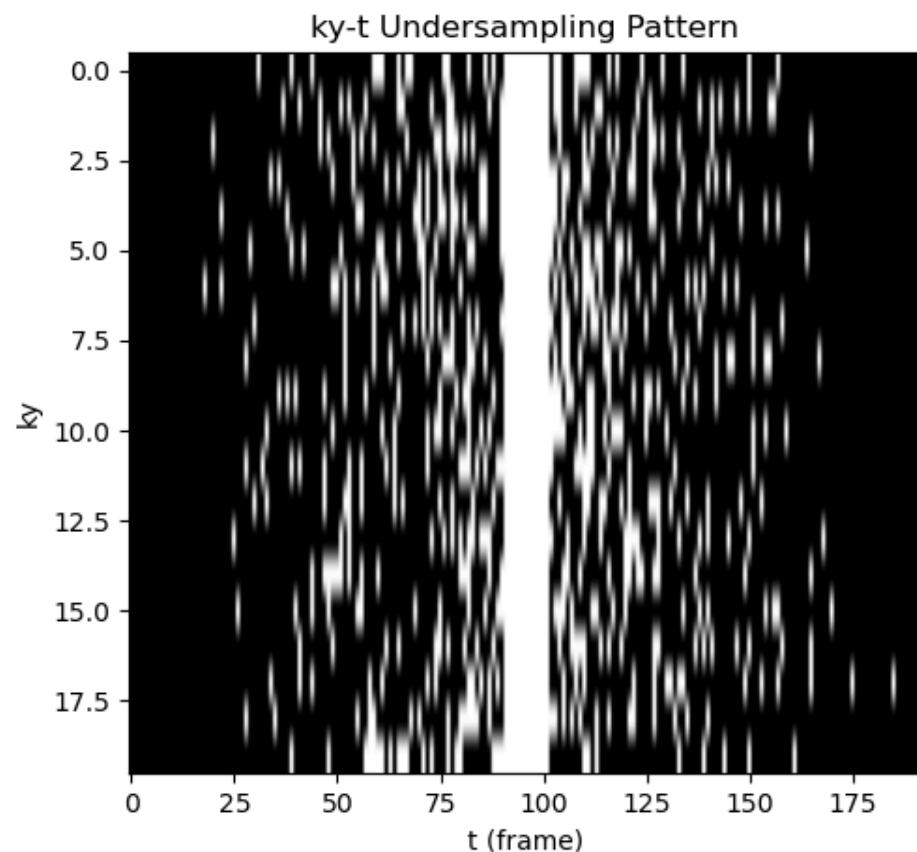
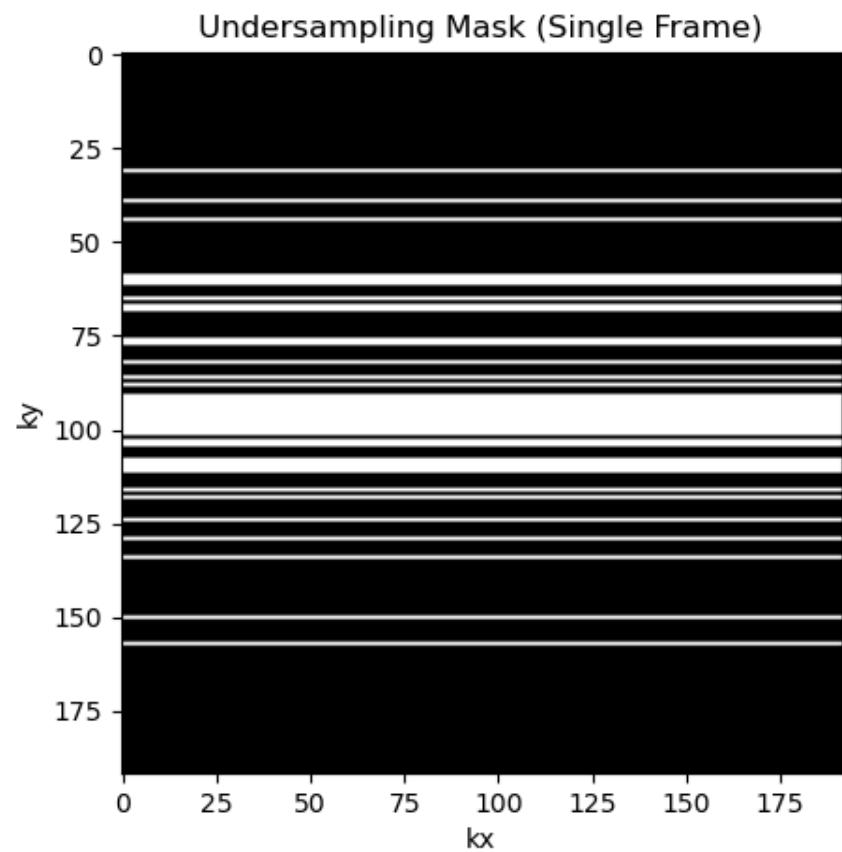
Member: 熊闻野 夏博扬 杨人一 吴家兴 杨丰敏

Overview

- **Goal:** Reconstruct high-quality dynamic MRI images from undersampled k-space data.
- **Challenge:** Undersampling introduces aliasing artifacts.
- **Approach:** Deep learning framework combining:
 - Dual 2D UNets (for real and imaginary components)
 - 3D ResNet (for temporal correlation)
- **Evaluation:** PSNR and SSIM metrics.

Data & Undersampling

- **Dataset:** `cine.npz` - Fully sampled cardiac cine MRI `[nsamples, nt, nx, ny]`.
- **Mask Generation:**
- Variable density random undersampling.
- Acceleration Factor (AF) = 5.
- 11 central k-space lines preserved per frame.
- Different masks for different frames.
- **Aliasing:** $b = F^{-1} \cdot U \cdot F \cdot m$



Aliased Images vs. Fully Sampled (1/3)

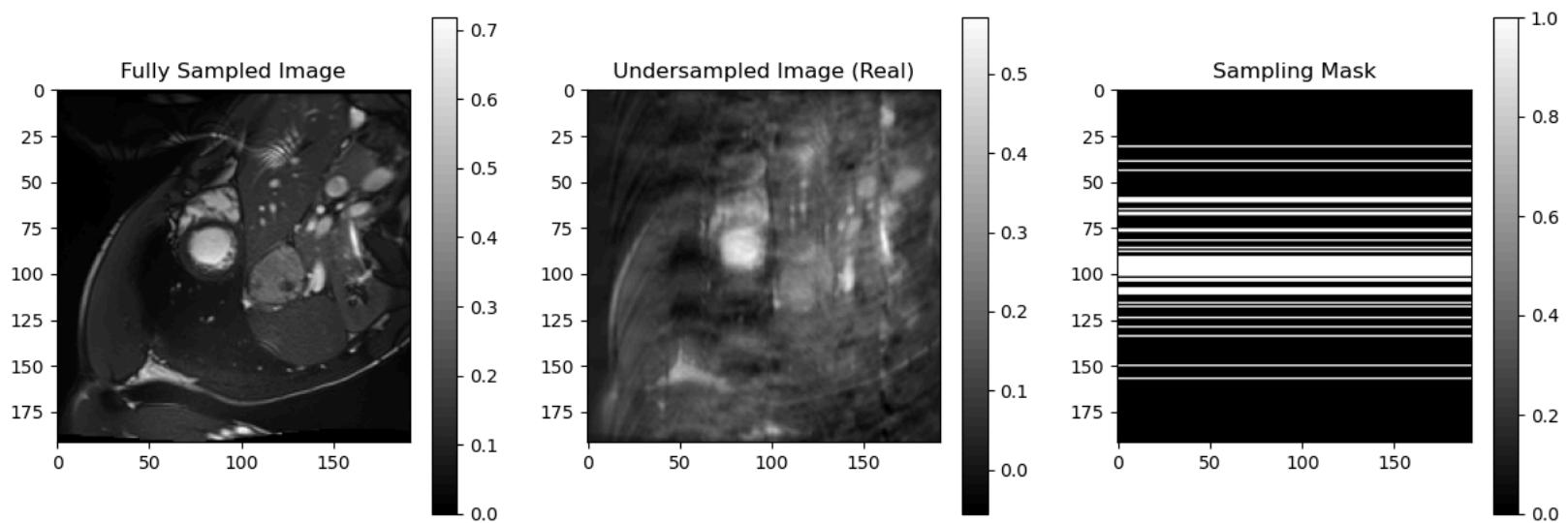


Fig: Fully sampled (left), Aliased (middle), Mask (right) - Frame 0

Aliased Images vs. Fully Sampled (2/3)

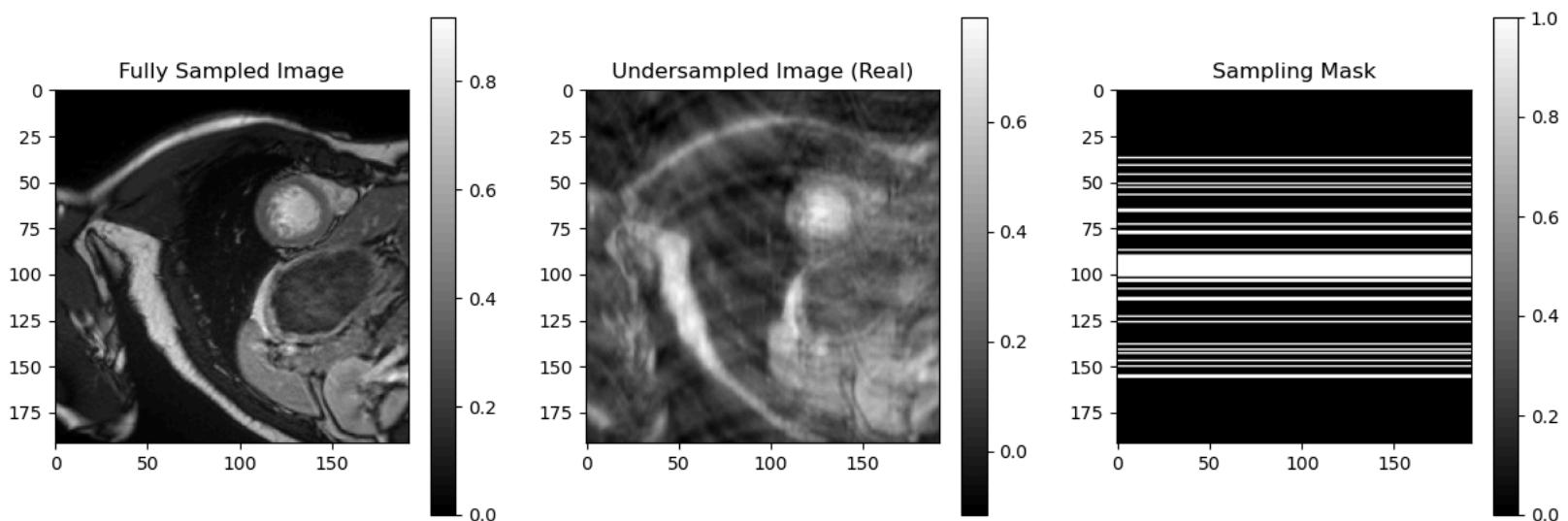


Fig: Fully sampled (left), Aliased (middle), Mask (right) - Frame 1

Aliased Images vs. Fully Sampled (3/3)

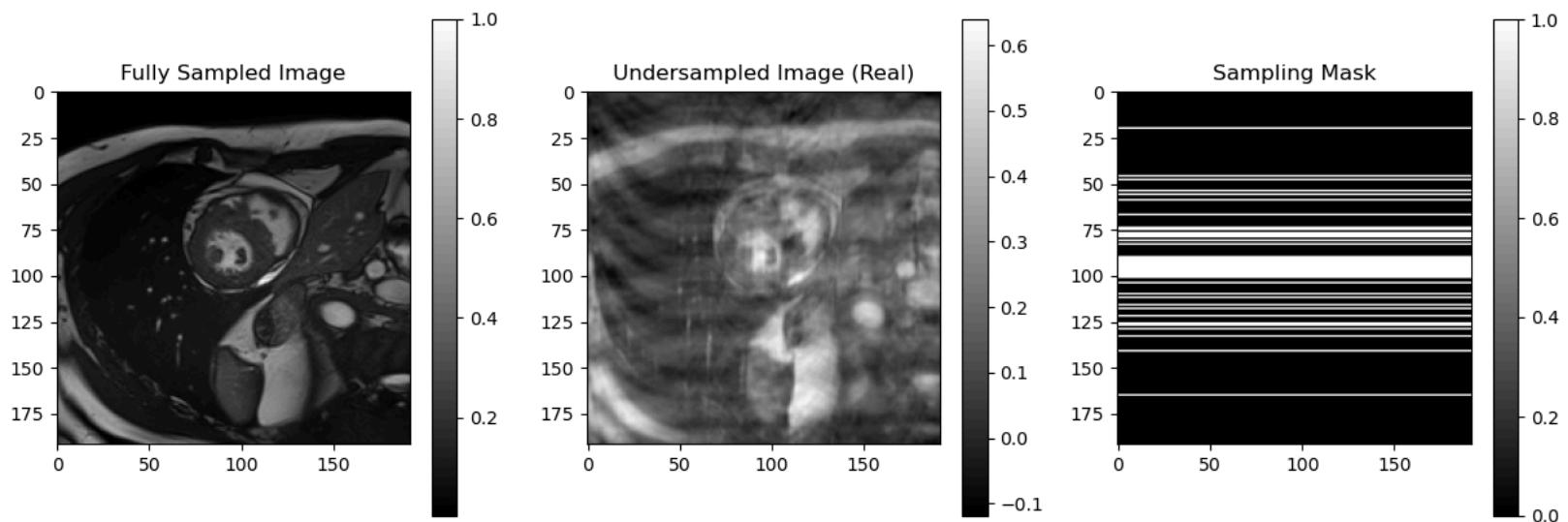
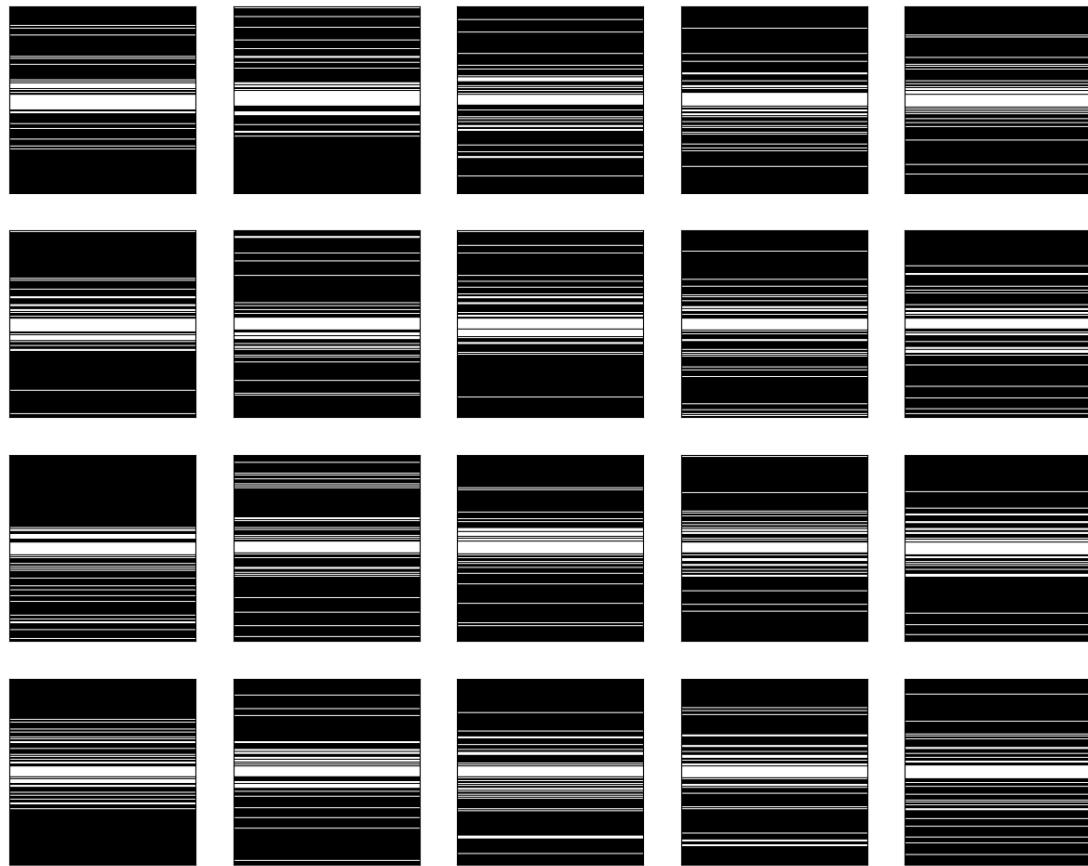


Fig: Fully sampled (left), Aliased (middle), Mask (right) - Frame 2



It is also clear to see that, for different dynamic frames, the undersampling masks are different.

Reconstruction Network: Dual UNet

- **Purpose:** Process real and imaginary parts separately.
- **Input:** Pseudo-complex images (real/imaginary as channels).
- **Features:**
 - Encoder-decoder with skip connections.
 - Attention mechanism (Channel & Spatial) in bottleneck.
 - Dropout ($p=0.3$).
 - LeakyReLU (negative_slope=0.1).
 - Weight Regularization.

Reconstruction Network: 3D ResNet

- **Purpose:** Integrate temporal information across frames.
- **Input:** Stacked outputs from the two UNets.
- **Features:**
 - 3D Convolutions.
 - Residual connections (BasicBlock).
 - Lightweight design (1 block/layer).
 - Final 1x1x1 convolution.

Creativity: Addressing Challenges in Network Design

- **Challenge 1: Pseudo-Complex Input:** Stacking dynamic images along the channel dimension created issues as real/imaginary parts weren't aligned.
- **Solution 1: Dual UNet Branches:** Split input into separate real and imaginary processing branches using two UNets, concatenating them later. Added attention in bottlenecks to enhance spatial/channel correlation capture.
- **Challenge 2: Capturing Temporal Correlation:** Standard 2D UNet structures don't effectively model changes over time.
- **Solution 2: 3D ResNet Integration:** Added a 3D ResNet structure after the UNets specifically to process and fuse information across the temporal dimension (frames).

Network Architecture Detail

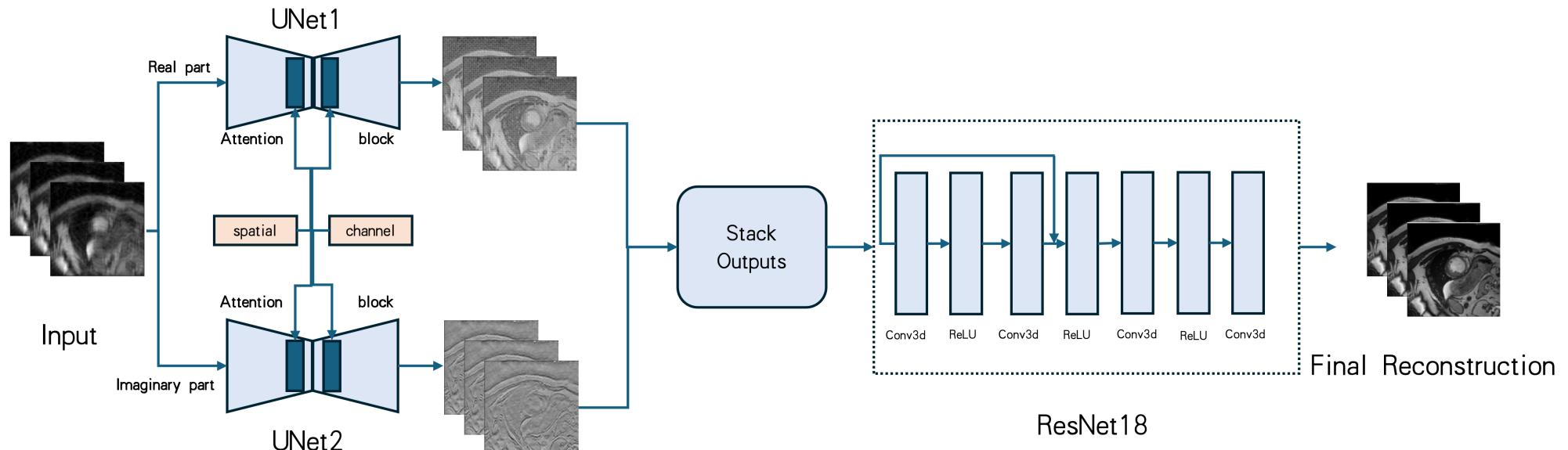
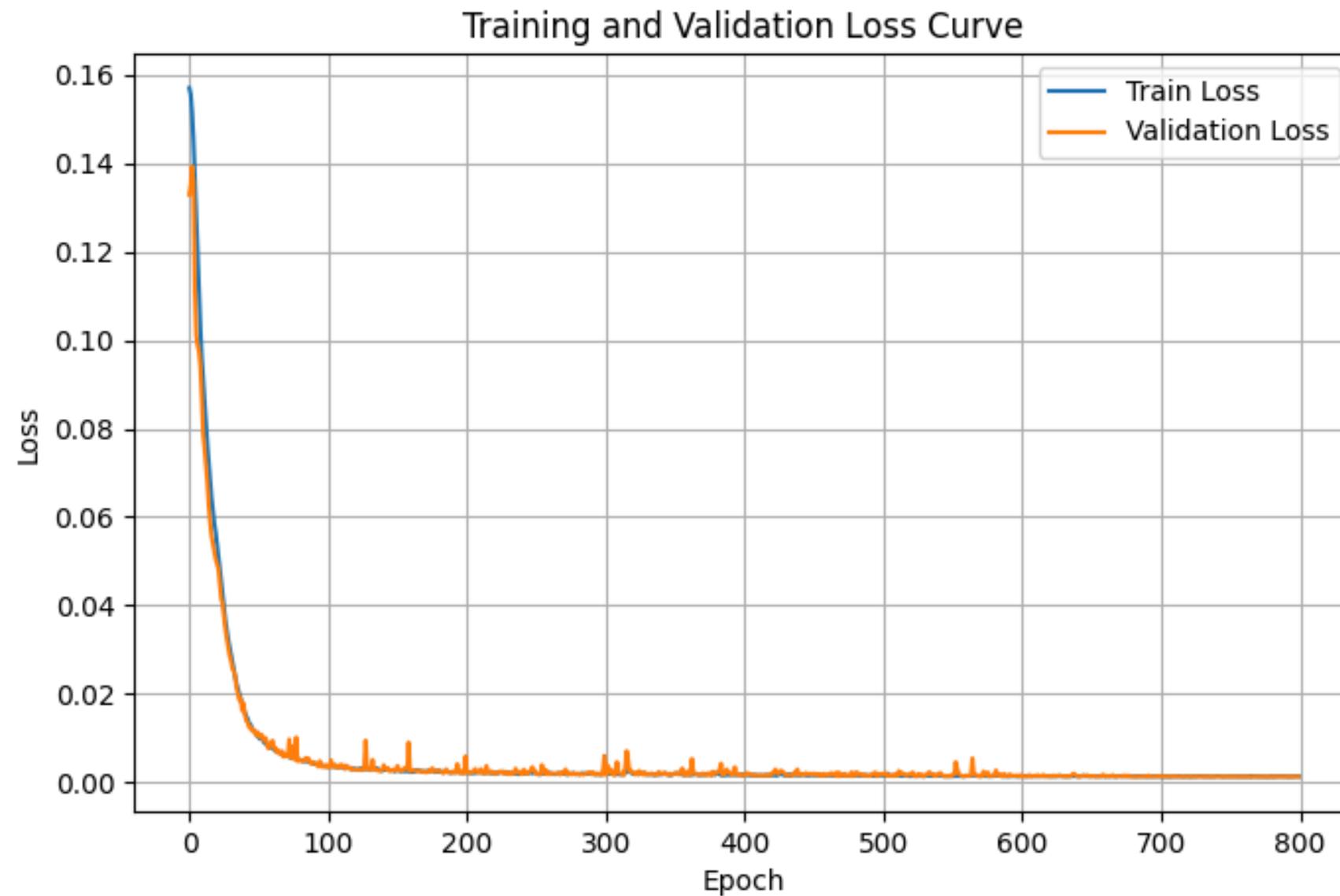


Fig: Detailed architecture showing dual UNet branches and 3D ResNet

Results: Main Model (L2 Loss)

- Metrics:
- Loss: mean = 0.00135 ± 0.00055
- PSNR: mean = 29.084 ± 1.932
- SSIM: mean = 0.844 ± 0.037
- Significant improvement over aliased images.



Reconstruction Examples (1/2)

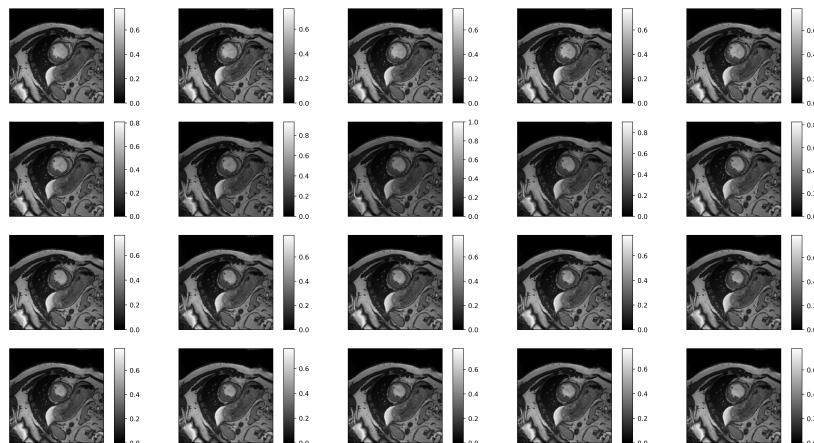


Fig: Fully Sampled (Ground Truth)

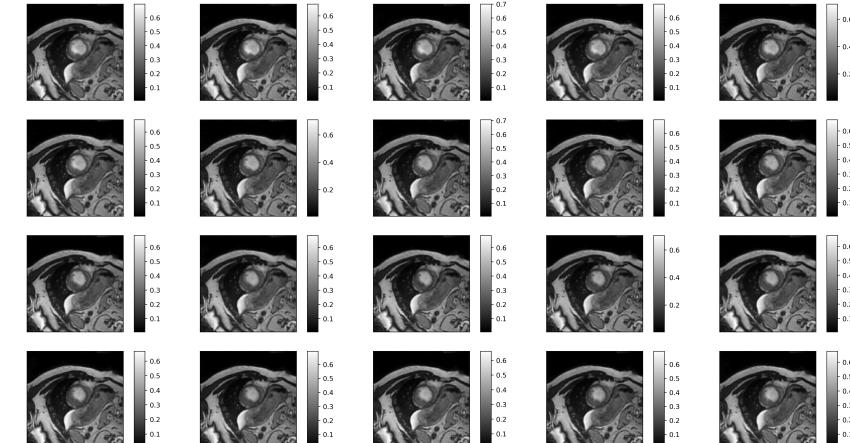


Fig: Reconstructed Image

Reconstruction Examples (2/2)

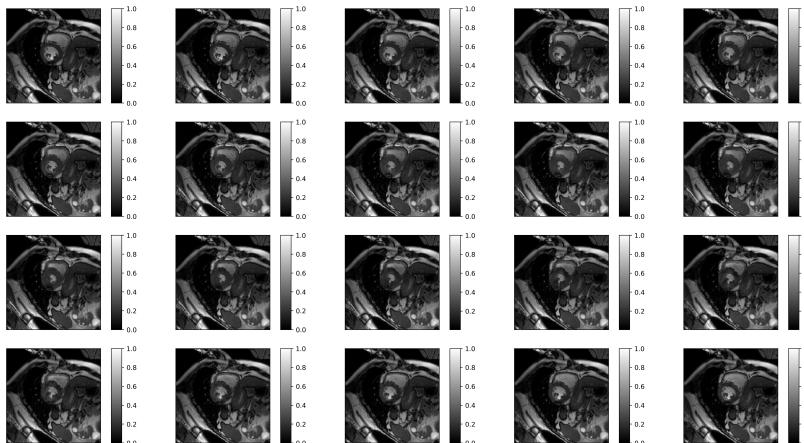


Fig: Fully Sampled (Ground Truth)

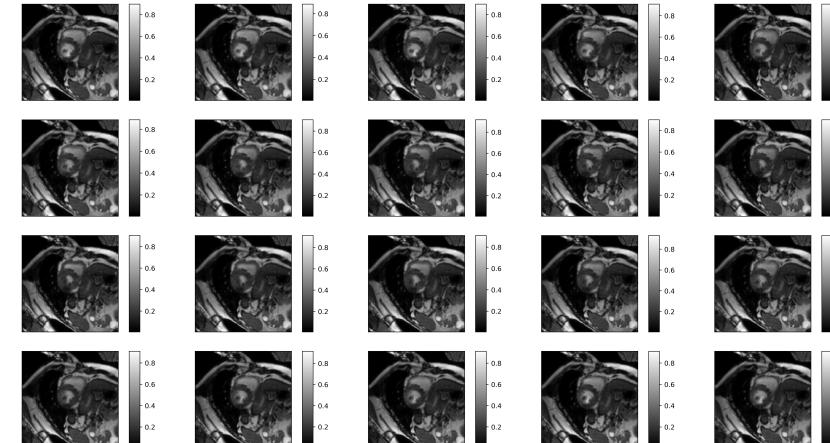
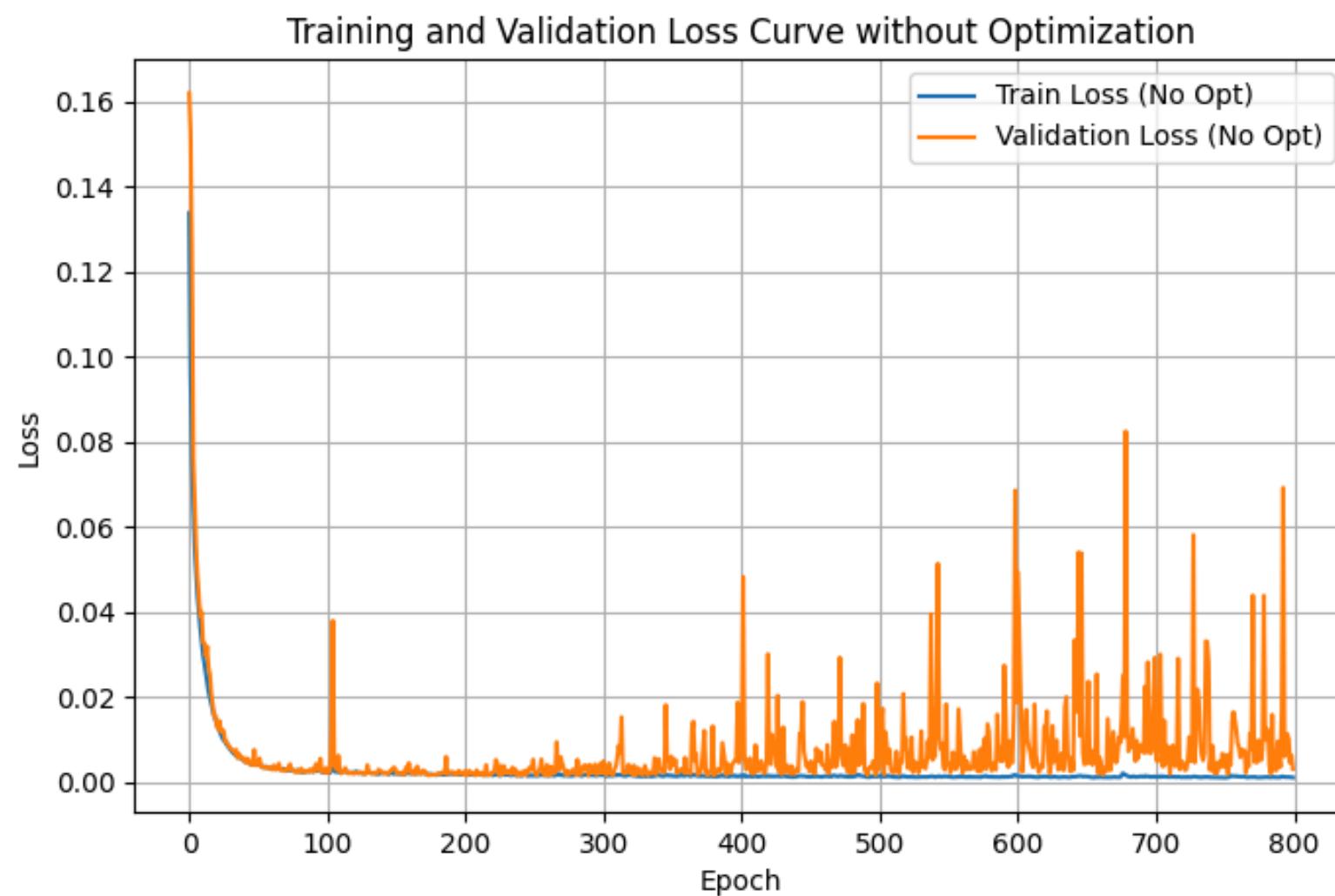


Fig: Reconstructed Image

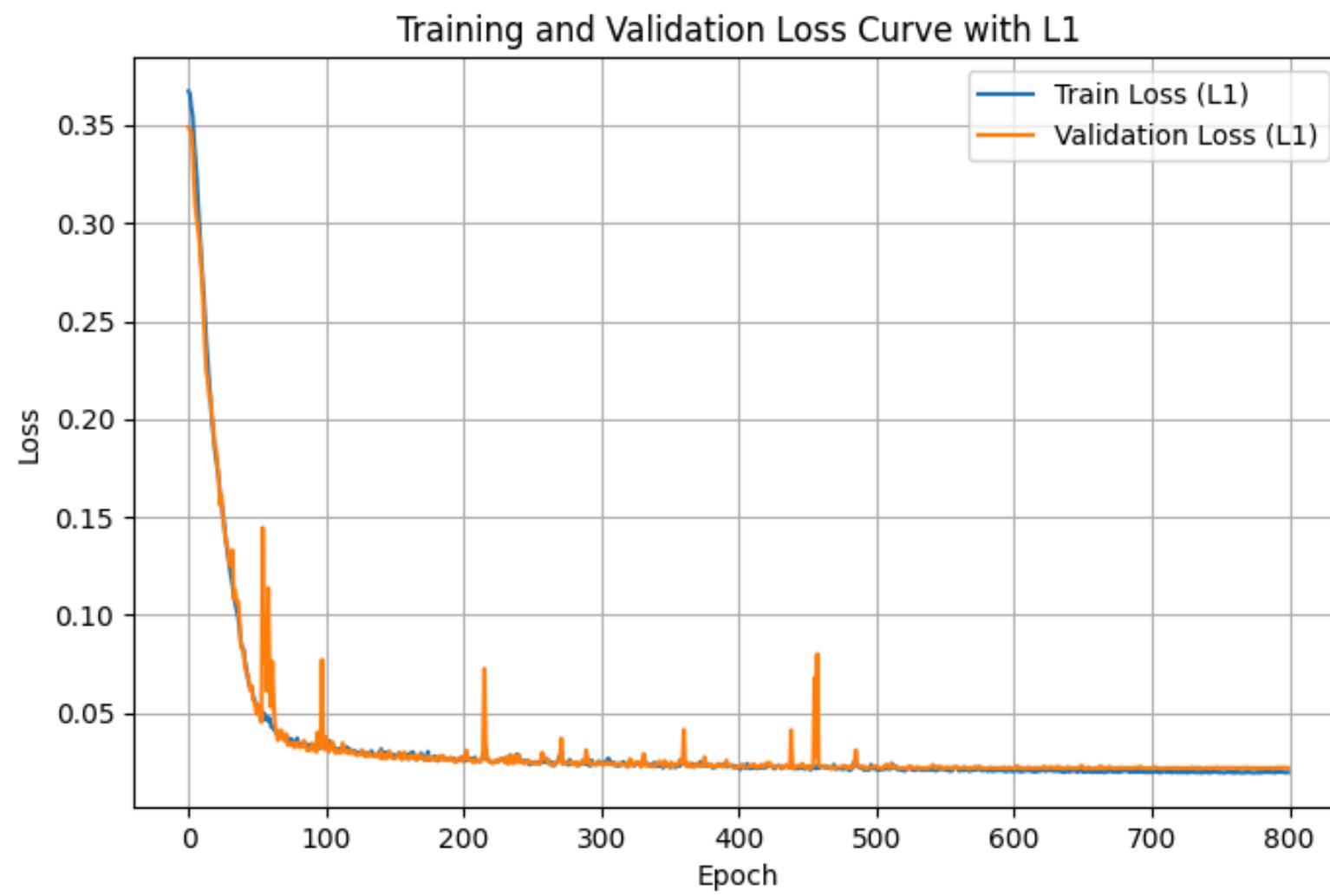
Ablation: Impact of Dropout & Dynamic LR

- **Model:** Trained without Dropout and with constant LR.
- **Results:**
- PSNR: 24.154 (vs. 29.084)
- SSIM: 0.743 (vs. 0.844)
- **Observation:** Clear signs of overfitting (validation loss). Dropout and dynamic LR are crucial for regularization and stable convergence.



Ablation: Impact of L1 vs. L2 Loss

- **Results (L1):** PSNR: 29.1511, SSIM: 0.8439
- **Results (L2):** PSNR: 29.0845, SSIM: 0.8443
- **PSNR vs. SSIM Trade-off:** L1 loss can lead to higher PSNR (pixel accuracy) but potentially lower SSIM (structural similarity) because it doesn't explicitly enforce structural consistency. In this specific case, SSIM was similar for both.
- **Observations:** Both loss functions yielded high-quality reconstructions. L2 loss resulted in much lower mean loss values and slightly better metric stability (lower std dev).
- **Recommendation:** Use L1/L2 if pixel recovery is the priority; consider structure-aware losses (e.g., L1+SSIM) if perceptual quality/structural fidelity is crucial. The original L2 model was retained for stability.



Exploration: Unrolled Denoising Network

- **Concept:** Cascade base network with data consistency layers.
- **Models:** 2 Cascades (C2), 3 Cascades (C3).
- **Training:** Increased memory/time significantly. Trained only 300 epochs.

Model	Epochs	GPU Mem	PSNR	SSIM
Original	800	~10GB	29.08	0.844
Cascade 2	300	18GB	28.87	0.834
Cascade 3	300	24GB	28.96	0.807

- **Observation:** Performance did not improve over original model, possibly due to limited training data/epochs or base network complexity.

Unrolled Network Loss (Cascade 3)

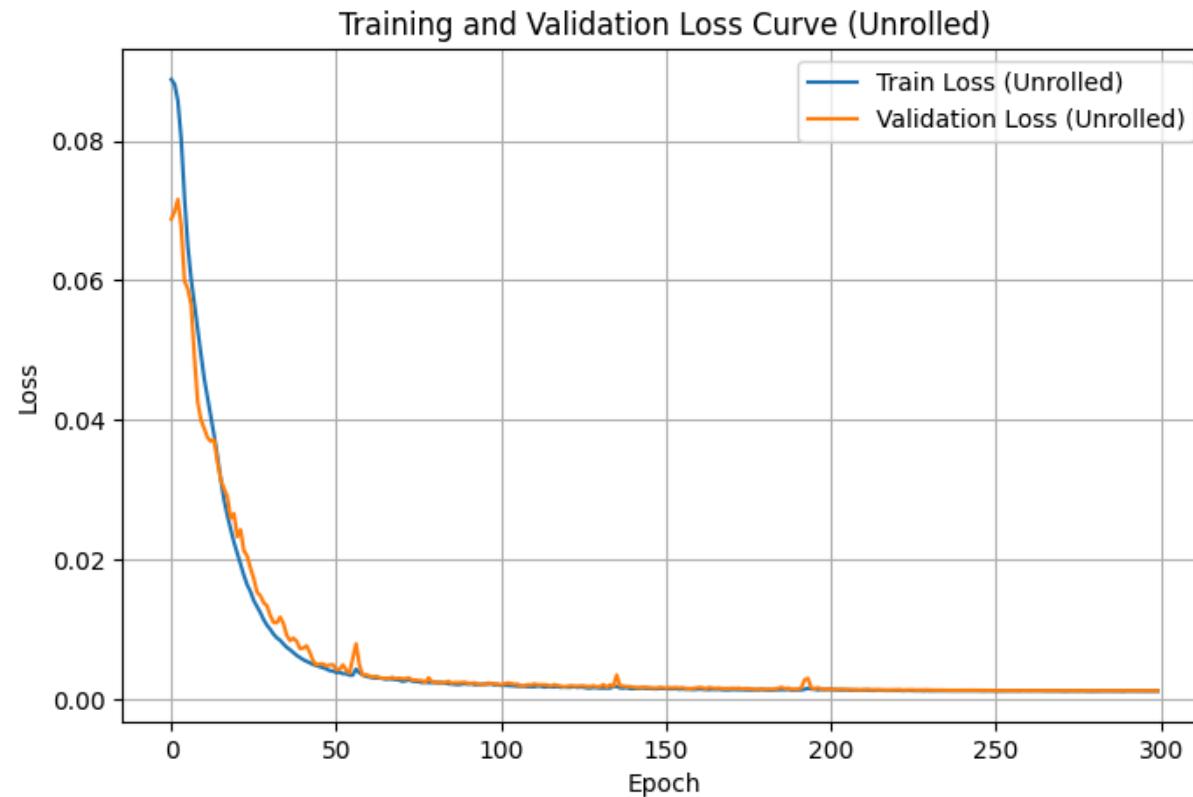


Fig: Loss Curves for 3-Cascade Unrolled Network (300 Epochs)

Conclusion

- Proposed Dual UNet + 3D ResNet architecture effectively reconstructs dynamic MRI from undersampled data (PSNR ~29.1, SSIM ~0.84).
- Dropout and dynamic learning rate are essential for optimal performance.
- L1 and L2 loss functions yield comparable results; L2 chosen for stability.
- Unrolled networks showed potential but require further investigation (more data, longer training).