Adversarial MRI Defense System Report

MRI Adversarial Purification Report

Generated on: 2025-04-07 18:41:04

Image: pgd_0.1_a7865442_adversarial.png

Error adding comparison image: cannot write mode RGBA as JPEG

Purification Metrics

PSNR: 7.52 dB

SSIM: 0.0001

Interpretation of Metrics:

PSNR (Peak Signal-to-Noise Ratio): 7.52 dB

This indicates moderate image quality after purification. There may be visible differences from the original.

SSIM (Structural Similarity Index): 0.0001

This indicates moderate structural preservation. There may be noticeable structural differences from the original.

Technical Details

Purification Model:

This purification was performed using a U-Net based deep learning model with attention mechanisms. The model was

trained in two phases:

1. Phase 1: Initial training with pixel-wise and perceptual losses to learn the mapping from adversarial to clean images.

2. Phase 2: GAN-based fine-tuning with a PatchGAN discriminator for improved sharpness and detail preservation.

Adversarial MRI Defense System Report

The model incorporates self-attention mechanisms, residual connections, and multi-scale feature fusion to effectively remove adversarial perturbations while preserving important diagnostic features.