

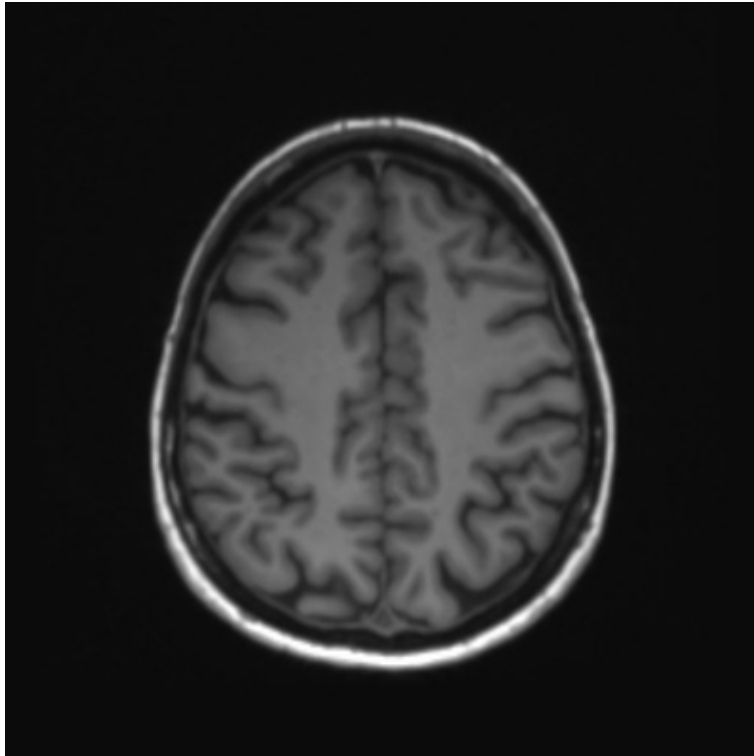
Adversarial MRI Defense System Report

MRI Adversarial Attack Classification Report

Generated on: 2025-04-19 23:14:45

Image: mri_healthy_1.jpeg

Original Image



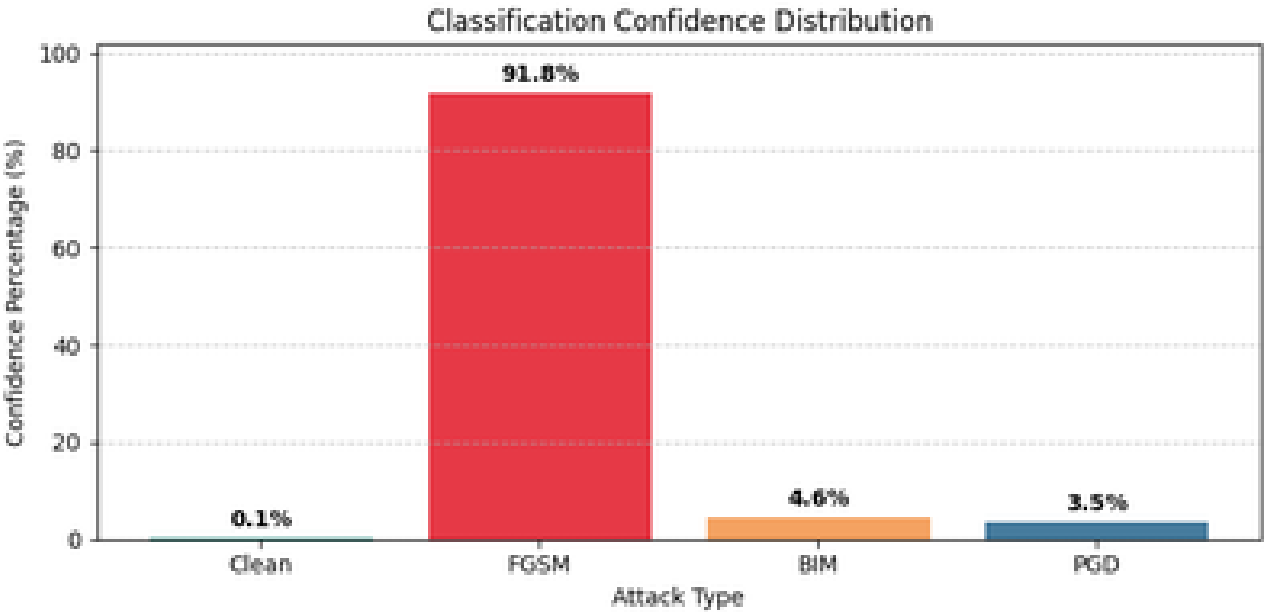
Classification Results

Predicted Class: FGSM

Confidence Scores:

Clean	0.12%
FGSM	91.78%
BIM	4.59%
PGD	3.51%

Adversarial MRI Defense System Report



Error generating model contribution chart: shape mismatch: objects cannot be broadcast to a single shape. Mismatch is between

Interpretation:

The image shows signs of Fast Gradient Sign Method attack, which is a one-step attack that perturbs the image in the direction of the gradient.

Recommendations:

Consider running the image through our purification system to remove adversarial perturbations.