Iterative Adversarial Training: Comprehensive Results **Model Robustness Attack Effectiveness** 1.0 1.000 0.975 8.0 0.950 Test Accuracy 0.900 0.875 Success Rate 0.6 0.850 0.2 0.825 0.800 0.0 1.5 Iteration 0.5 1.5 2.0 2.5 0.0 0.5 2.0 2.5 0.0 1.0 3.0 1.0 3.0 Iteration **Attack Difficulty Perturbation Efficiency** 3.70 0.28150 0.28125 3.65 0.28100 0.28075 0.28050 0.28025 Avg Perturbations
52
69
69 0.28000 3.50 0.27975 1.5 0.5 1.0 2.0 2.5 0.5 1.0 1.5 2.0 2.5 0.0 3.0 0.0 3.0

Iteration

Iteration