

Defense Implemented: Randomization

Randomization Approach: Random Resizing

Metric	Clean	No Defense Attack	w/ Defense Attack
Loss	0.0312	4.88	0.00104
Accuracy	98%	6%	100%

Example Misclassifications:

Number of misclassified images for Clean: 1
Attack: fgsm_cw_attack
Dataset: MNIST
Training Epochs: 10
Trained Clean Images: 54210
Test Images: 140
Accuracy: 0.98
Precision: 0.99
Recall: 0.99
F1-score: 0.99
ROC AUC Score: 1.00

Misclassifications:
0 -> 8: 1

Number of misclassified images for No Defense Attack: 60
Attack: fgsm_cw_attack
Dataset: MNIST
Training Epochs: 10
Adversarial Training Images: 27105
Test Images: 140
Accuracy: 0.06
Precision: 0.12
Recall: 0.10
F1-score: 0.11
ROC AUC Score: 1.00

Misclassifications:
0 -> 5: 4
0 -> 6: 34
0 -> 2: 9
0 -> 9: 8
0 -> 7: 2
0 -> 8: 2
0 -> 4: 1

No misclassified images for stage: w/ Defense Attack
Attack: fgsm_cw_attack
Dataset: MNIST
Training Epochs: 10
Retrained Clean and Adversarial Images: 81315
Test Images: 140
Accuracy: 1.00
Precision: 1.00
Recall: 1.00
F1-score: 1.00
ROC AUC score is not defined for a single class.
