

Defense Implemented: Randomization

Randomization Approach: Random Resizing

Metric	Clean	No Defense Attack	w/ Defense Attack
Loss	0.00128	0.0119	0.0751
Accuracy	100%	98%	95%

Example Misclassifications:

No misclassified images for stage: Clean  
Attack: cw\_pgd\_attack  
Dataset: MNIST  
Training Epochs: 10  
Trained Clean Images: 54210  
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.

Number of misclassified images for No Defense Attack: 1  
Attack: cw\_pgd\_attack  
Dataset: MNIST  
Training Epochs: 10  
Adversarial Training Images: 13939  
Test Images: 140  
Accuracy: 0.98  
Precision: 0.99  
Recall: 0.99  
F1-score: 0.99  
ROC AUC Score: 1.00

Misclassifications:  
0 -> 6: 1

Number of misclassified images for w/ Defense Attack: 3  
Attack: cw\_pgd\_attack  
Dataset: MNIST  
Training Epochs: 10  
Retrained Clean and Adversarial Images: 68149  
Test Images: 140

Accuracy: 0.95  
Precision: 0.98  
Recall: 0.97  
F1-score: 0.97  
ROC AUC Score: 1.00

---

Misclassifications:

0 -> 6: 2

0 -> 9: 1