Defense Implemented: Input Transformation

Input Transformation Approach: Combined Input Transformation

Metric	Clean	No Defense Attack	w/ Defense Attack
Loss	0.000423	0.0299	0.001
Accuracy	100%	98%	100%

Example Misclassifications:

No misclassified images for stage: Clean
Attack: pgd_bim_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: pgd_bim_attack
Dataset: MNIST
Training Epochs: 10
Adversarial Training Images: 13598
Test Images: 140
Accuracy: 0.98
Precision: 0.99
Recall: 0.99
F1-score: 0.99
ROC AUC Score: 1.00

Misclassifications:
0 -> 2: 1

No misclassified images for stage: w/ Defense Attack
Attack: pgd_bim_attack
Dataset: MNIST
Training Epochs: 10
Retrained Clean and Adversarial Images: 67808
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.
