```
Defense Implemented: Input Transformation

Input Transformation Approach: Adversarial Logit Pairing

+----------+---------+--------------------+---------------------+
| Metric   | Clean   | No Defense Attack   | w/ Defense Attack   |
+==========+=========+====================+=====================+
| Loss     | 0.0124  | 0.158              | 0.0232              |
+----------+---------+--------------------+---------------------+
| Accuracy | 100%    | 92%                | 98%                 |
+----------+---------+--------------------+---------------------+

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: 5
Attack: pgd_bim_attack
Dataset: MNIST
Training Epochs: 10
Adversarial Training Images: 9376
Test Images: 140
Accuracy: 0.92
Precision: 0.96
Recall: 0.96
F1-score: 0.96
ROC AUC Score: 1.00
------------------------------------------------------

Misclassifications:
0 -> 6: 3
0 -> 9: 1
0 -> 5: 1

------------------------------------------------------
Number of misclassified images for w/ Defense Attack: 1
Attack: pgd_bim_attack
Dataset: MNIST
Training Epochs: 10
```

```
Retrained Clean and Adversarial Images: 63586
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
```