

Metric	Clean	No Defense Attack	w/ Defense Attack
Loss	0.00278	0.134	0.0116
Accuracy	100%	92%	100%

#### Example Misclassifications:

```

-----
No misclassified images for stage: Clean
Attack: cw_pgd_attack
Dataset: MNIST
Training Epochs: 10
Trained Clean Images: 64
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: cw_pgd_attack
Dataset: MNIST
Training Epochs: 10
Adversarial Training Images: 60
Test Images: 140
Accuracy: 0.92
Precision: 0.96
Recall: 0.96
F1-score: 0.96
ROC AUC Score: 1.00
-----

```

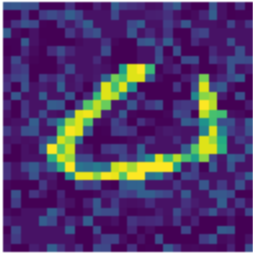
#### Misclassifications:

```

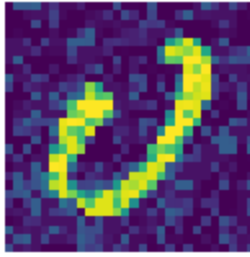
0 -> 4: 1
0 -> 7: 1
0 -> 2: 1
0 -> 9: 1
0 -> 6: 1

```

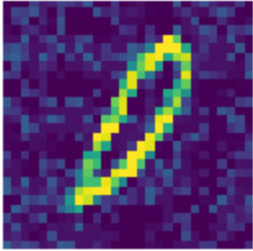
0 -&gt; 4



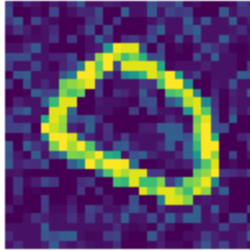
0 -&gt; 7



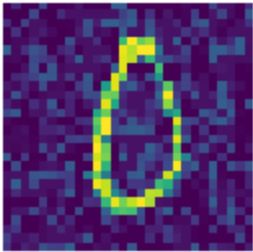
0 -&gt; 2



0 -&gt; 9



0 -&gt; 6



-----  
No misclassified images for stage: w/ Defense Attack

Attack: cw\_pgd\_attack

Dataset: MNIST

Training Epochs: 10

Retrained Clean and Adversarial Images: 124

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.