Metric	Clean	No Defense Attack	++ w/ Defense Attack
Loss	0.00546	4.1	0.00546
Accuracy	100%	17%	100%

Example Misclassifications:

No misclassified images for stage: Clean

Attack: fgsm_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: 53

Attack: fgsm_pgd_attack

Dataset: MNIST Training Epochs: 10

Adversarial Training Images: 60

Test Images: 140
Accuracy: 0.17
Precision: 0.31
Recall: 0.28
F1-score: 0.29

Misclassifications:

ROC AUC Score: 1.00

0 -> 6: 31

0 -> 7: 4

0 -> 5: 5

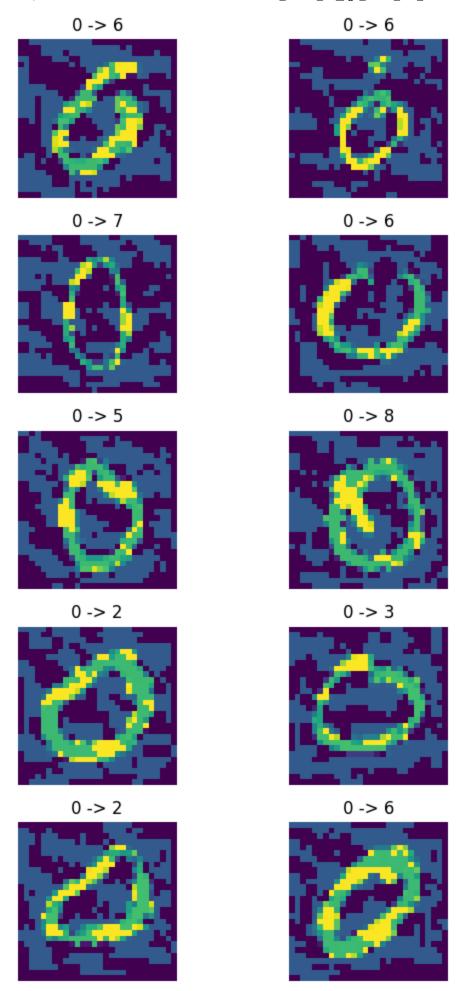
0 -> 8: 3

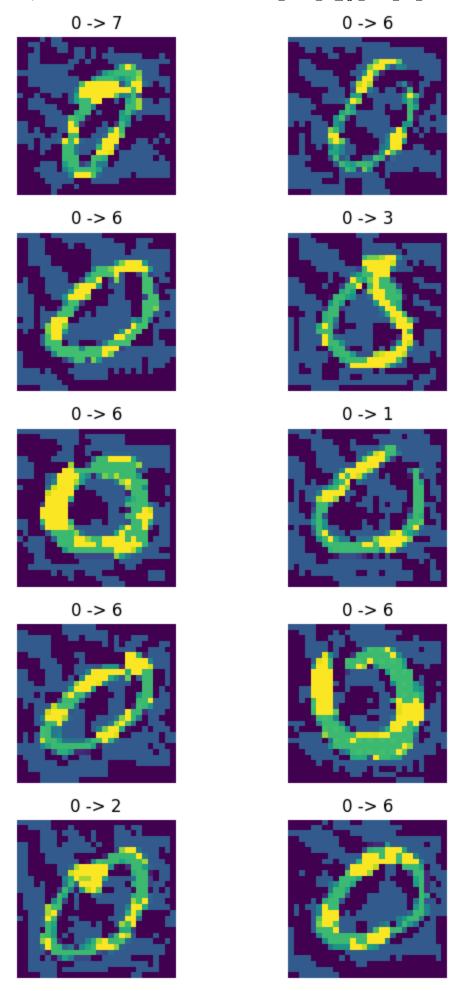
0 -> 2: 6

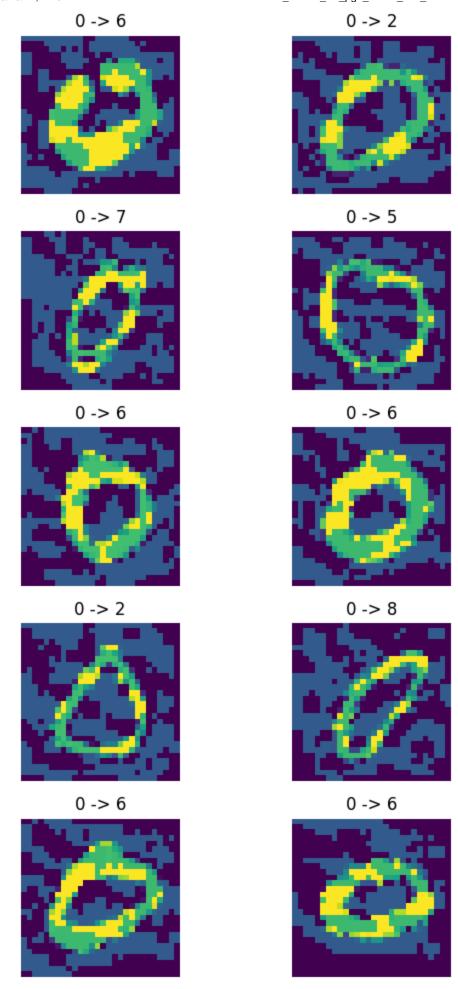
0 -> 3: 2

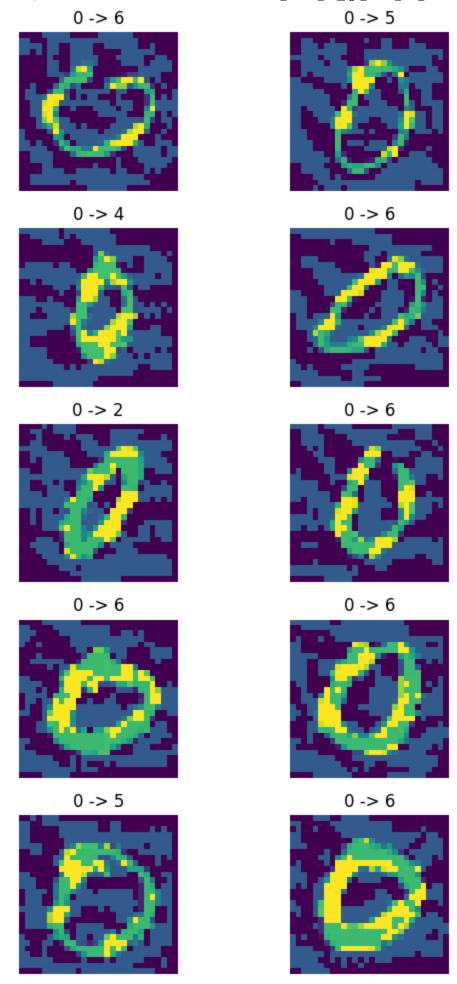
0 -> 1: 1

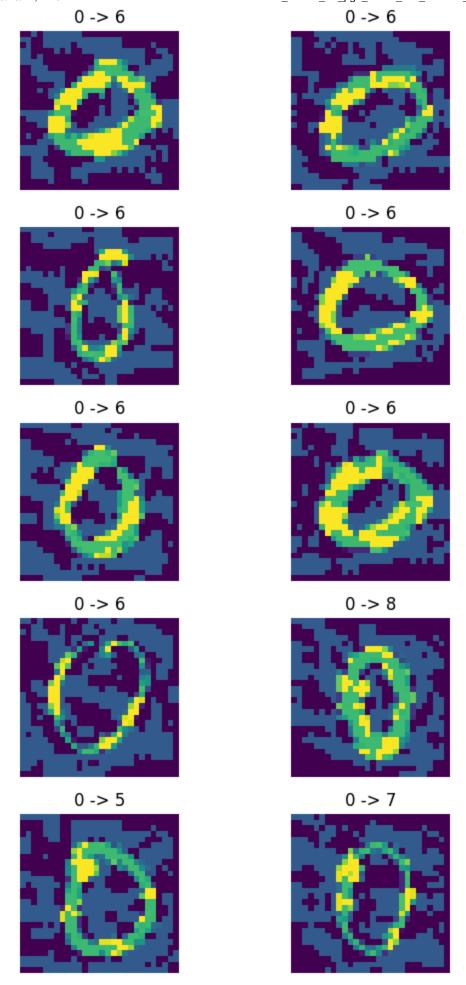
0 -> 4: 1



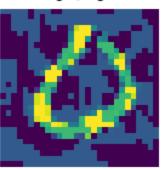




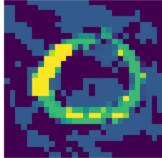




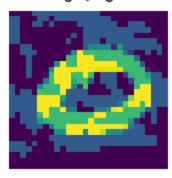
0 -> 6



0 -> 6



0 -> 6



No misclassified images for stage: w/ Defense Attack

Attack: fgsm_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.