

| Metric | Clean | No Defense Attack | w/ Defense Attack |
|----------|--------|-------------------|-------------------|
| Loss | 0.0778 | 4.19 | 0.0186 |
| Accuracy | 98% | 8% | 100% |

Example Misclassifications:

Number of misclassified images for Clean: 1

Attack: fgsm_deepfool_attack

Dataset: MNIST

Training Epochs: 10

Trained Clean Images: 64

Test Images: 140

Accuracy: 0.98

Precision: 0.99

Recall: 0.99

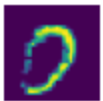
F1-score: 0.99

ROC AUC Score: 1.00

Misclassifications:

0 -> 7: 1

0 -> 7



Number of misclassified images for No Defense Attack: 59

Attack: fgsm_deepfool_attack

Dataset: MNIST

Training Epochs: 10

Adversarial Training Images: 60

Test Images: 140

Accuracy: 0.08

Precision: 0.14

Recall: 0.12

F1-score: 0.13

ROC AUC Score: 1.00

Misclassifications:

0 -> 7: 8

0 -> 8: 6

0 -> 6: 18

0 -> 9: 8

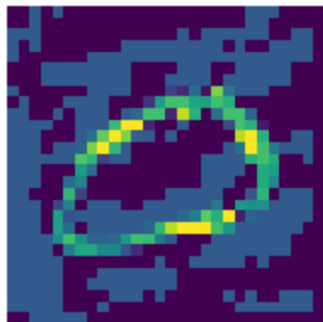
0 -> 2: 10

0 -> 4: 4

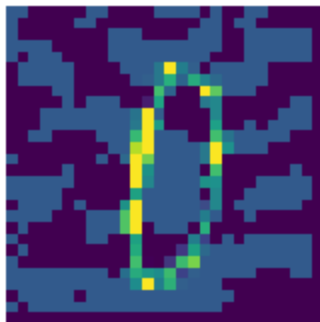
0 -> 5: 4

0 -> 1: 1

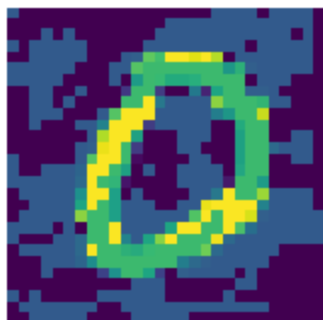
0 -> 7



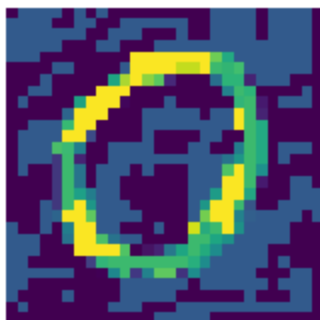
0 -> 8



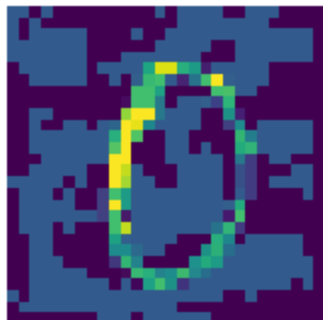
0 -> 6



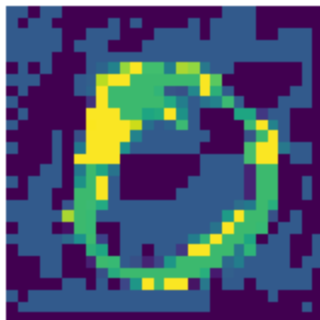
0 -> 9



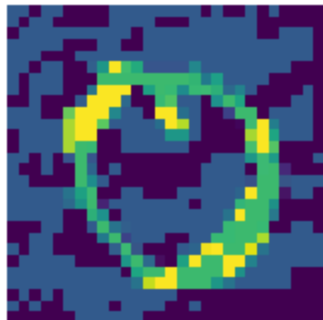
0 -> 6



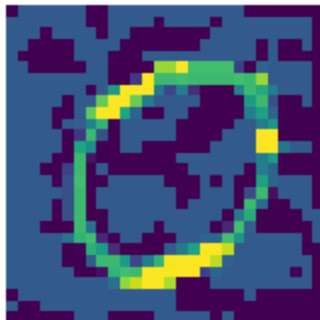
0 -> 7



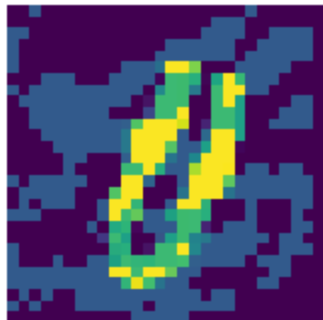
0 -> 2



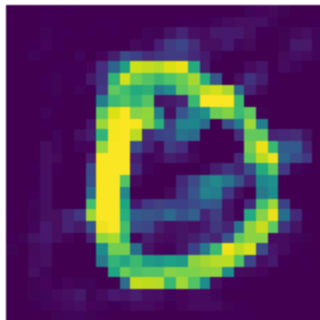
0 -> 2



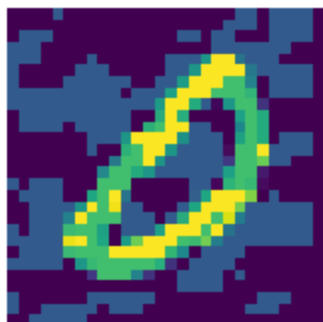
0 -> 4



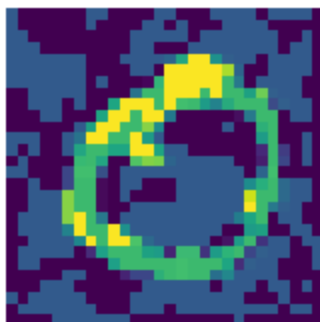
0 -> 6



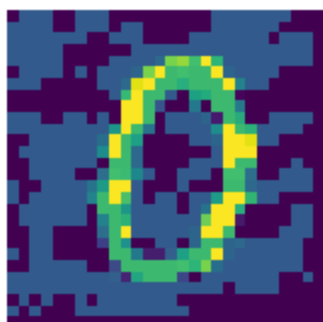
0 -> 6



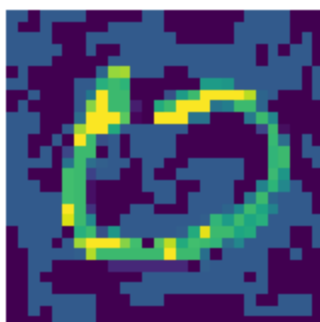
0 -> 9



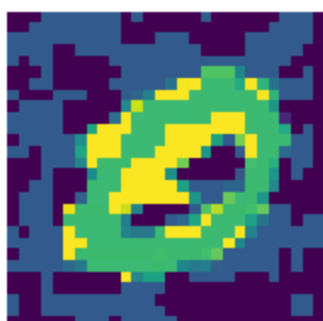
0 -> 7



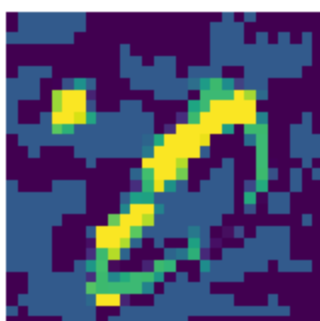
0 -> 5



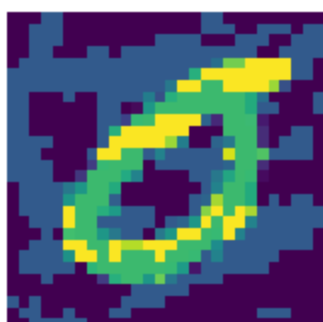
0 -> 8



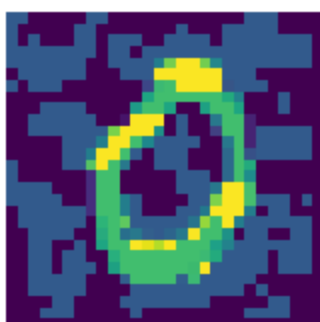
0 -> 8



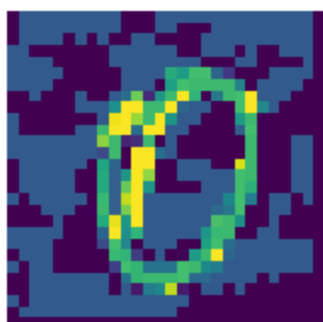
0 -> 4



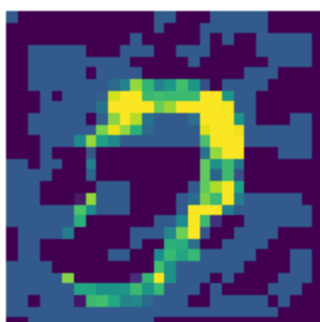
0 -> 6



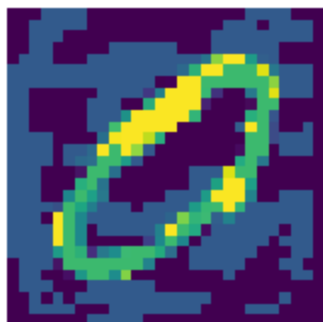
0 -> 8



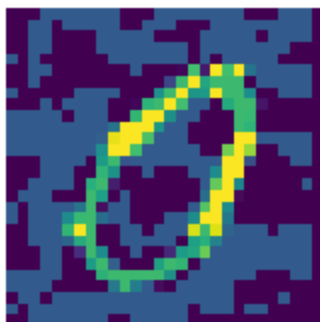
0 -> 7



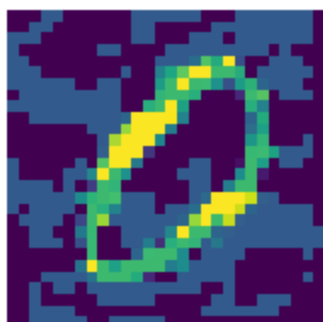
0 -> 2



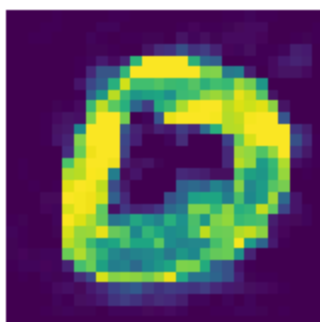
0 -> 7



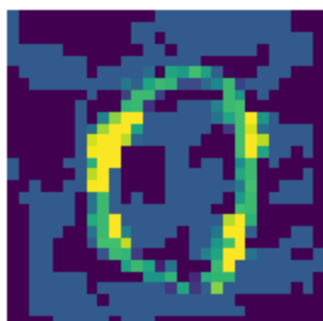
0 -> 6



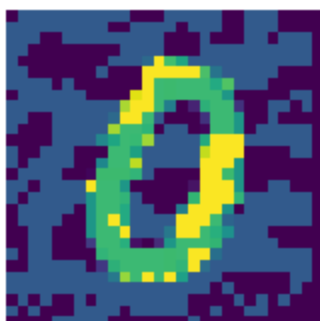
0 -> 6



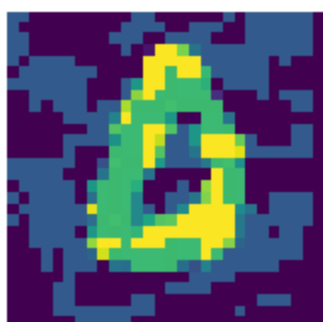
0 -> 4



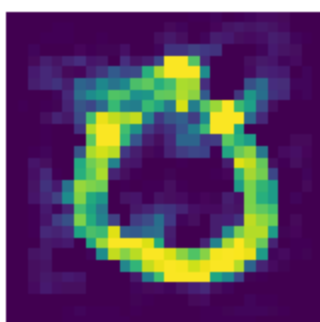
0 -> 7



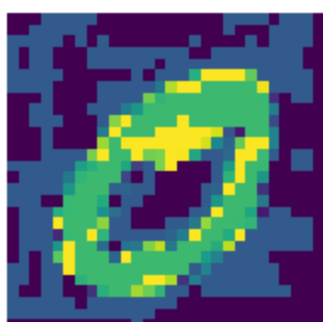
0 -> 2



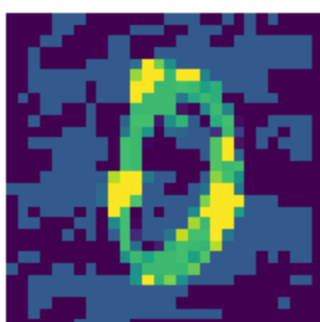
0 -> 5



0 -> 2



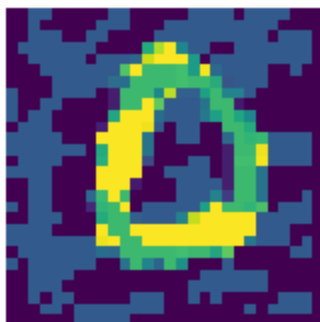
0 -> 4



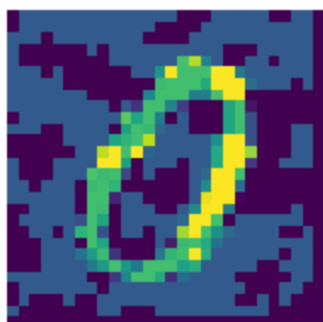
0 -> 6



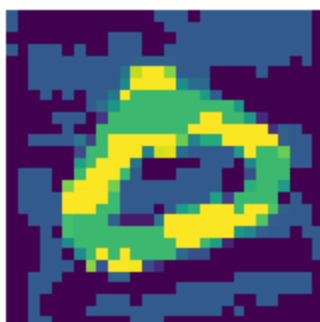
0 -> 6



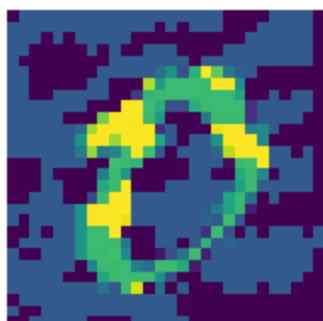
0 -> 7



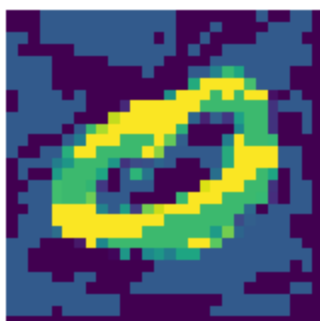
0 -> 6



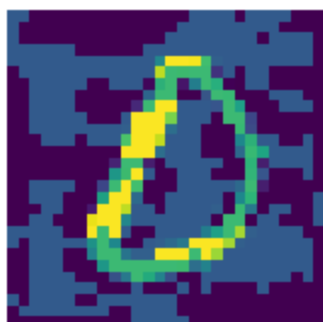
0 -> 2



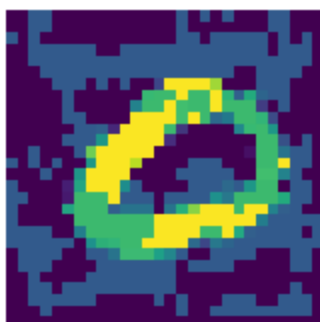
0 -> 9



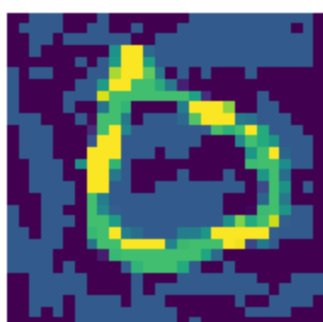
0 -> 2



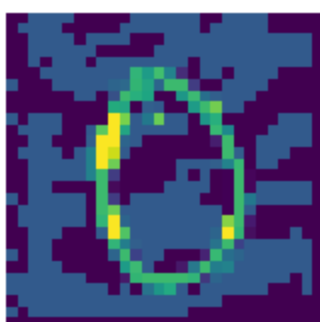
0 -> 6



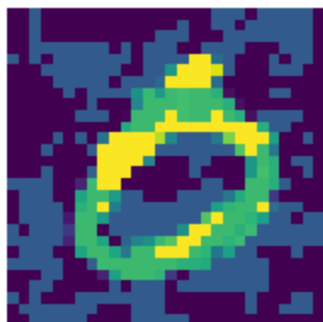
0 -> 6



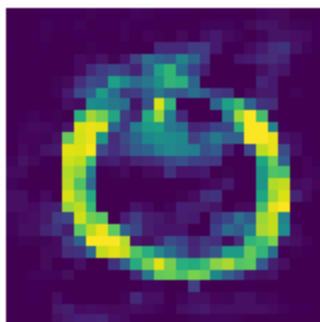
0 -> 6



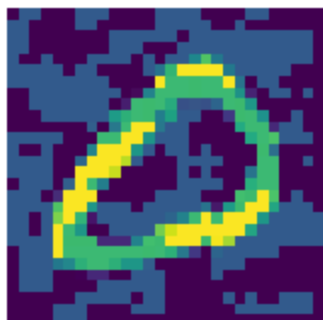
0 -> 6



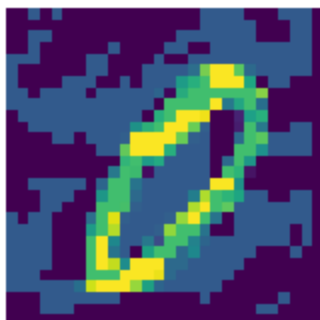
0 -> 5



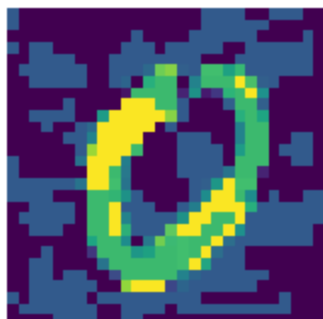
0 -> 2



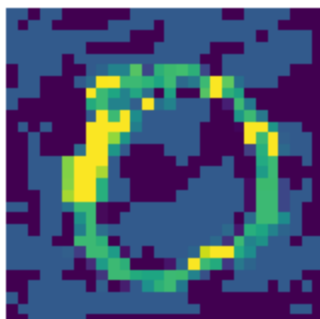
0 -> 8



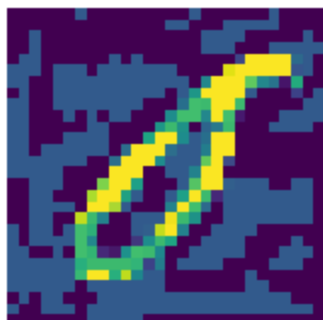
0 -> 6



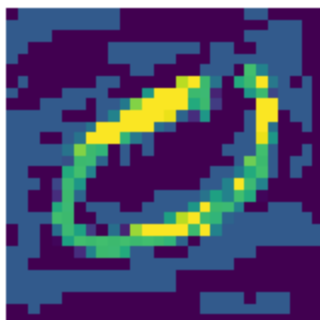
0 -> 6



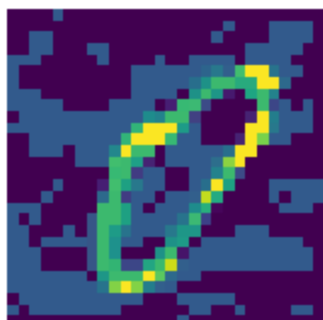
0 -> 1



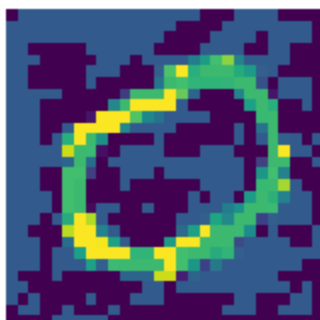
0 -> 2



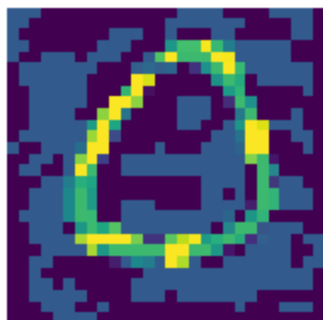
0 -> 8



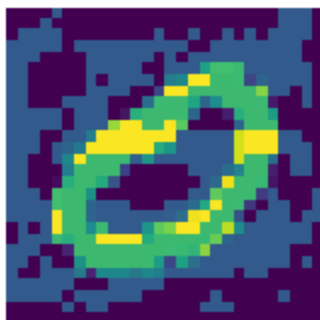
0 -> 9



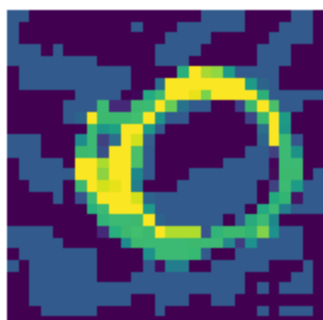
0 -> 9



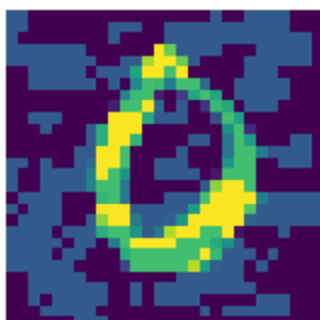
0 -> 2



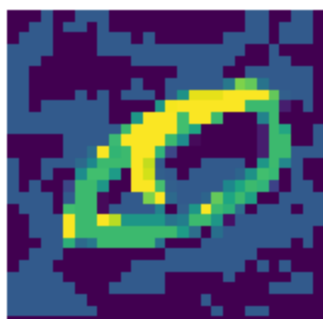
0 -> 6



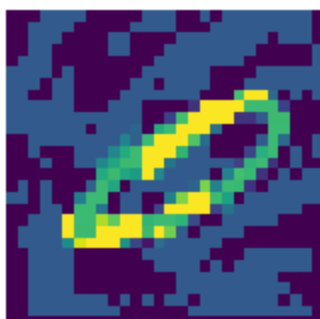
0 -> 6



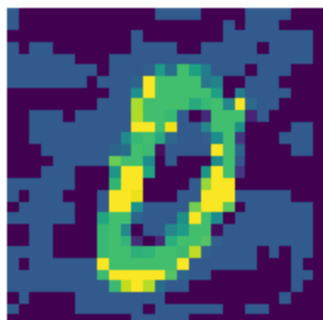
0 -> 9



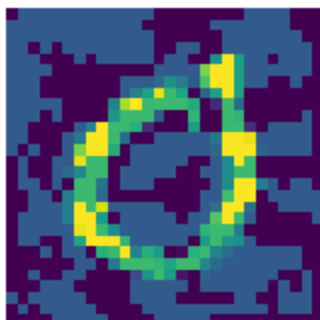
0 -> 5



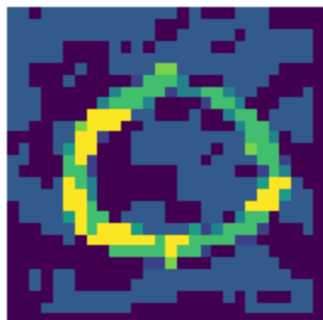
0 -> 7



0 -> 9



0 -> 9



```
-----  
No misclassified images for stage: w/ Defense Attack  
Attack: fgsm_deepfool_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.  
-----
```