Metric	Clean	No Defense Attack	w/ Defense Attack
Loss	0.00649	3.52	0.00649
Accuracy	100%	27%	100%

## Example Misclassifications:

-----

No misclassified images for stage: Clean

Attack: fgsm\_cw\_attack

Dataset: EMNIST Training Epochs: 10 Trained Clean Images: 64

Test Images: 625 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: 47

Attack: fgsm\_cw\_attack

Dataset: EMNIST Training Epochs: 10

Adversarial Training Images: 60

Test Images: 625 Accuracy: 0.27 Precision: 0.78 Recall: 0.29 F1-score: 0.42 ROC AUC Score: 1.00

------

## Misclassifications:

0 -> 2: 18

0 -> 6: 3

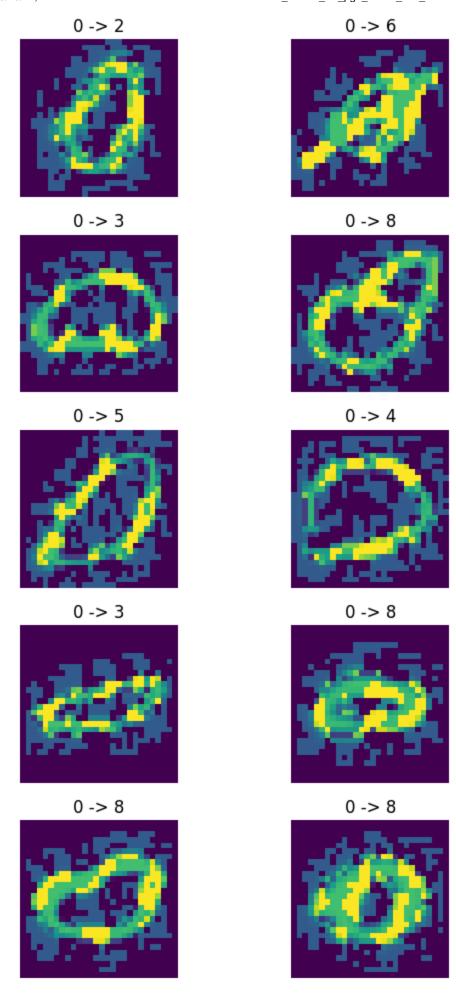
0 -> 3: 7

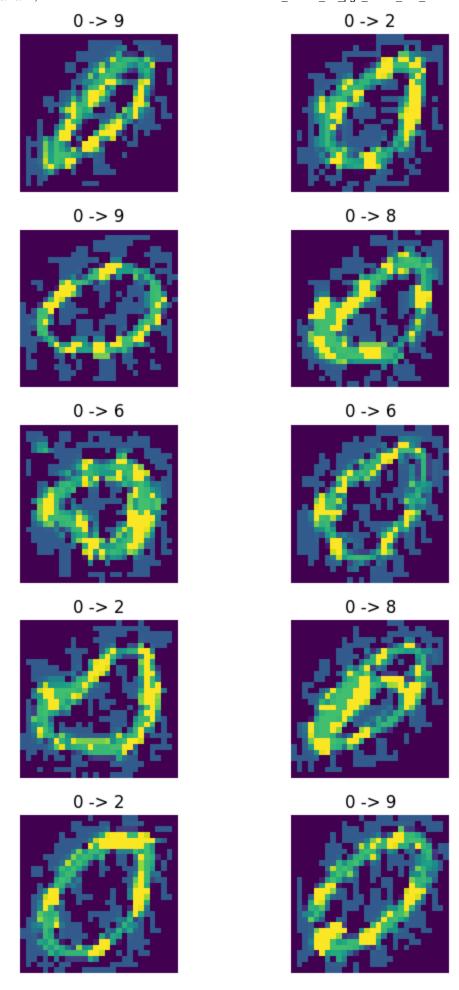
0 -> 8: 11

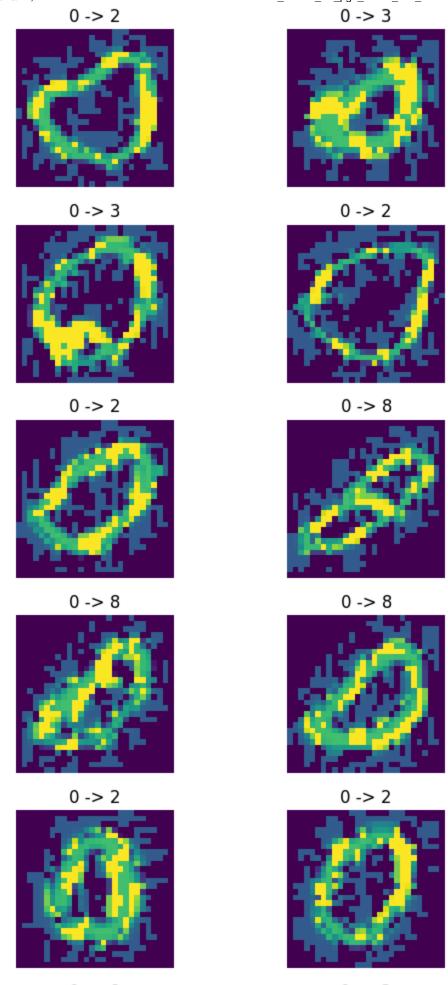
0 -> 5: 2

0 -> 4: 2

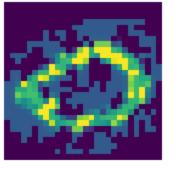
0 -> 9: 4



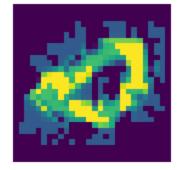




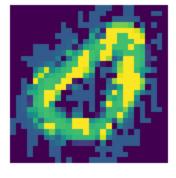




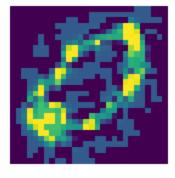




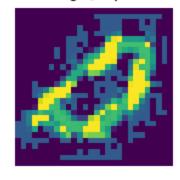
0 -> 2



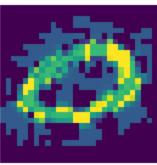
0 -> 2



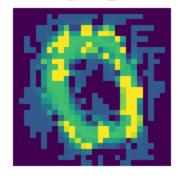
0 -> 4



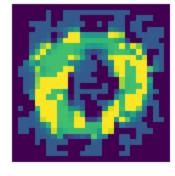
 $0 \rightarrow 2$ 



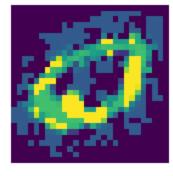
0 -> 2



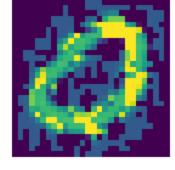
0 -> 3



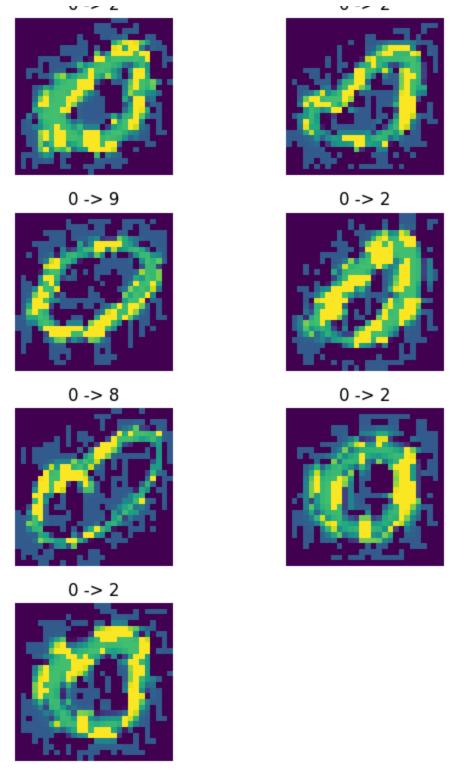
0 -> 3



0 -> 2



0-> 2



No misclassified images for stage: w/ Defense Attack

Attack: fgsm\_cw\_attack

Dataset: EMNIST Training Epochs: 10

Retrained Clean and Adversarial Images: 124

Test Images: 625 Accuracy: 1.00 Precision: 1.00 Recall: 1.00 F1-score: 1.00

ROC AUC score is not defined for a single class.

\_\_\_\_\_