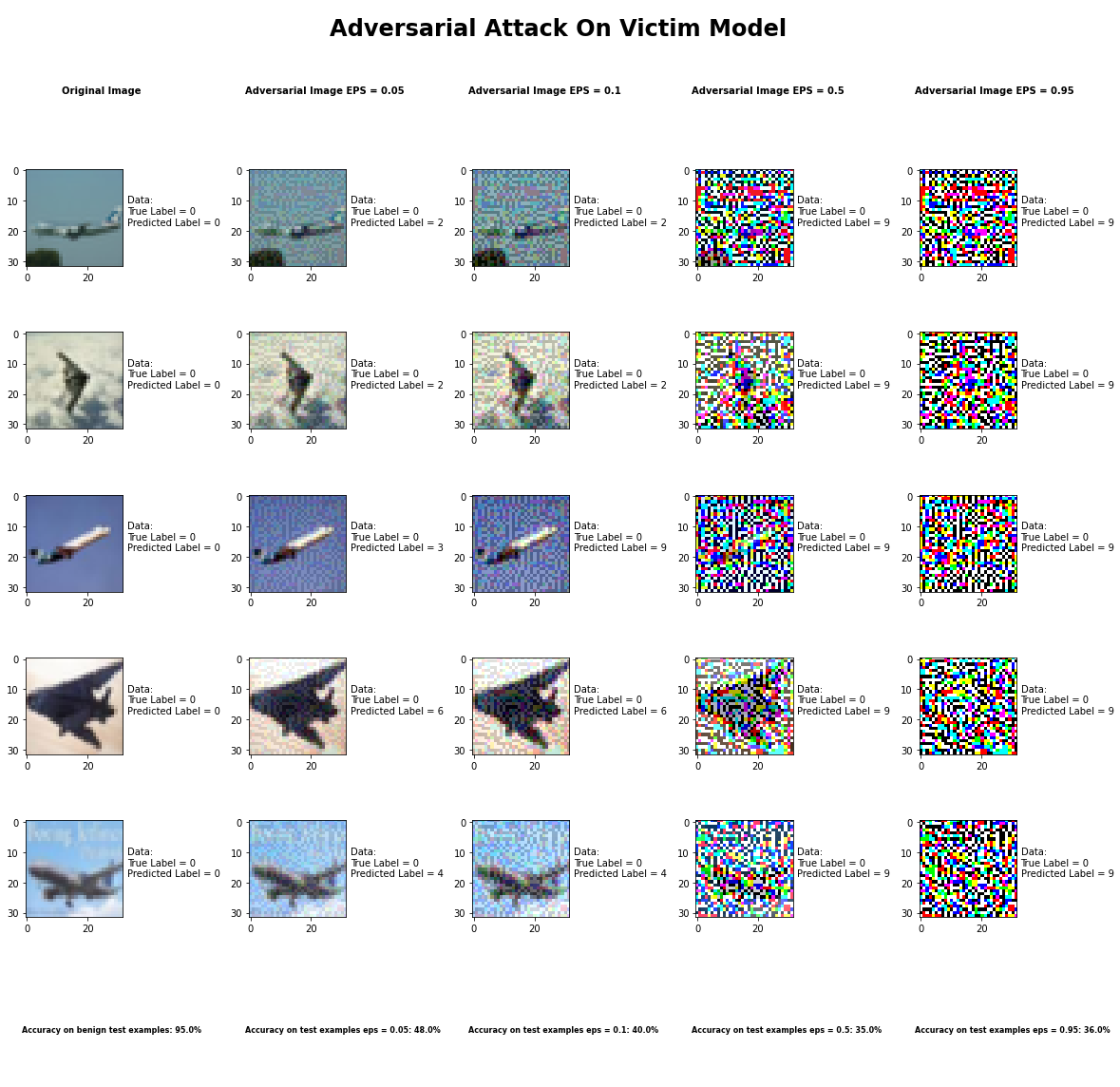
Fast Gradient Sign Method

|  |  |
| --- | --- |
| Average Accuracy Over 10 Runs | |
| Benign Samples | 94.52% |
| Epsilon = 0.05 Samples | 48.4% |
| Epsilon = 0.10 Samples | 39.9% |
| Epsilon = 0.5 Samples | 35.3% |
| Epsilon = 0.95 Samples | 35.3% |

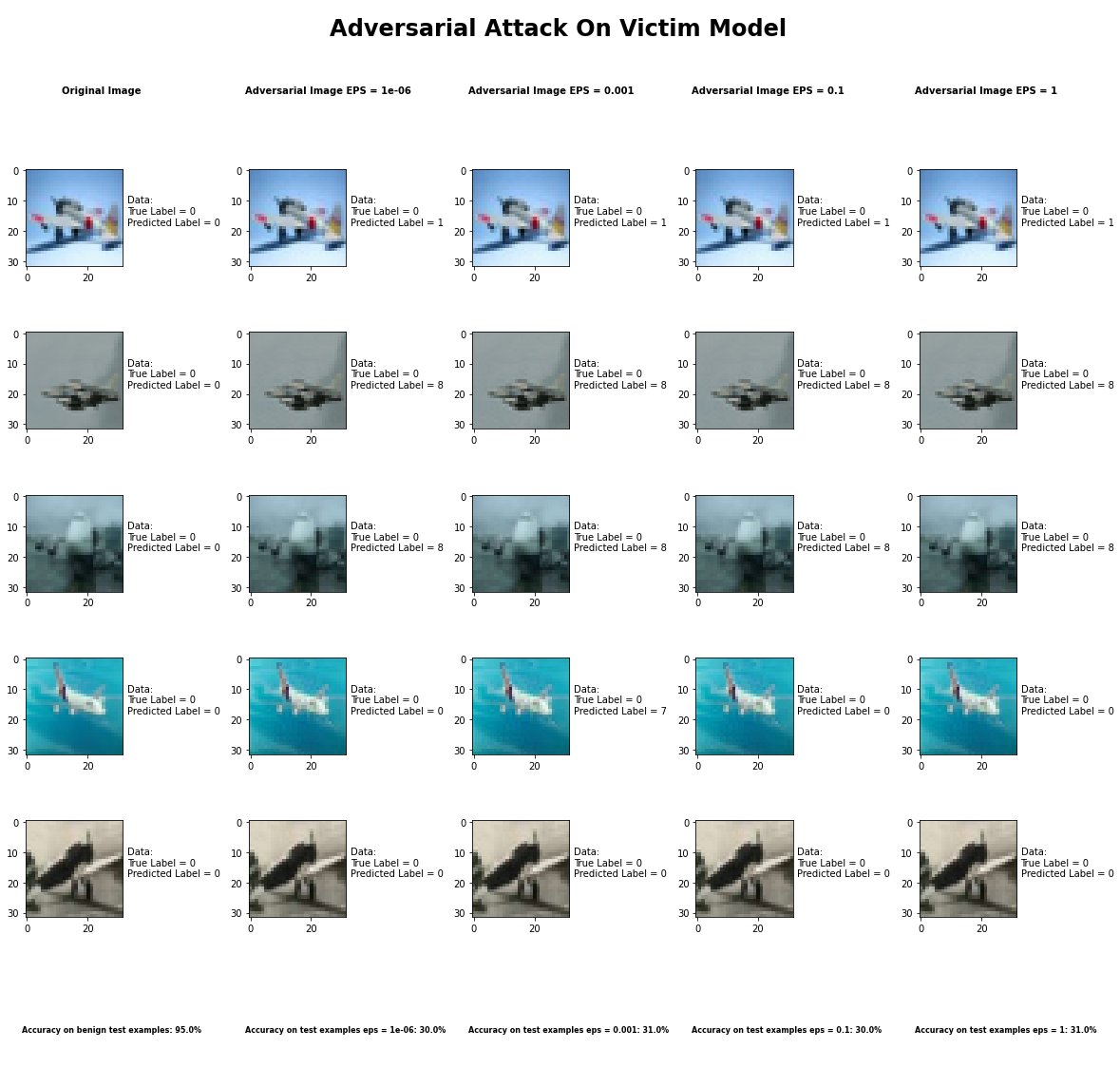
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Average Accuracy per Class Over 10 Runs | | | | | | | | | | |
|  | Airplane | Car | Bird | Cat | Deer | Dog | Frog | Horse | Ship | Truck |
| Benign Samples | 98% | 96% | 95% | 86% | 96% | 84% | 96% | 93% | 99% | 97% |
| Epsilon = 0.05 Samples | 32% | 60% | 68% | 42% | 43% | 31% | 60% | 33% | 58% | 57% |
| Epsilon = 0.10 Samples | 30% | 40% | 70% | 34% | 28% | 23% | 68% | 24% | 45% | 37% |
| Epsilon = 0.5 Samples | 33% | 40% | 34% | 37% | 20% | 23% | 26% | 24% | 45% | 71% |
| Epsilon = 0.95 Samples | 30% | 35% | 34% | 32% | 20% | 23% | 26% | 23% | 45% | 87% |



DeepFool Method

|  |  |
| --- | --- |
| Average Accuracy Over 10 Runs | |
| Benign Samples | 94.52% |
| Epsilon = 1e-6 Samples | 30.2% |
| Epsilon = 1e-3 Samples | 31.3% |
| Epsilon = 1e-1 Samples | 30.1% |
| Epsilon = 1 Samples | 31.0% |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Average Accuracy per Class Over 10 Runs | | | | | | | | | | |
|  | Airplane | Car | Bird | Cat | Deer | Dog | Frog | Horse | Ship | Truck |
| Benign Samples | 95% | 92% | 96% | 86% | 95% | 88% | 99% | 93% | 94% | 98% |
| Epsilon = 1e-6 Samples | 42% | 45% | 27% | 29% | 37% | 20% | 26% | 21% | 32% | 26% |
| Epsilon = 1e-3 Samples | 46% | 51% | 31% | 26% | 33% | 22% | 24% | 25% | 30% | 24% |
| Epsilon = 1e-1 Samples | 46% | 48% | 24% | 22% | 38% | 23% | 29% | 24% | 28% | 21% |
| Epsilon = 1  Samples | 46% | 41% | 28% | 24% | 35% | 22% | 25% | 26% | 32% | 18% |

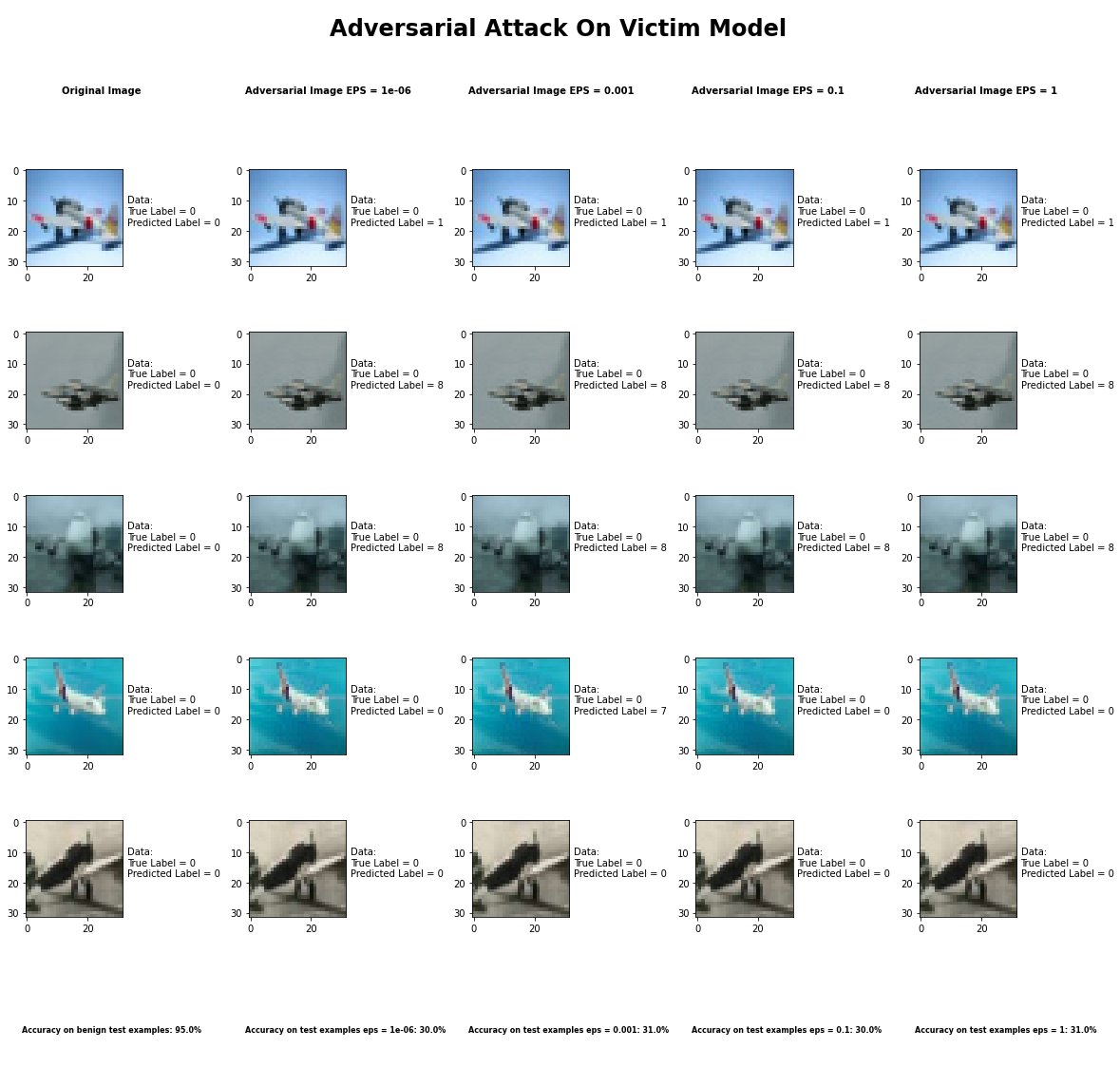


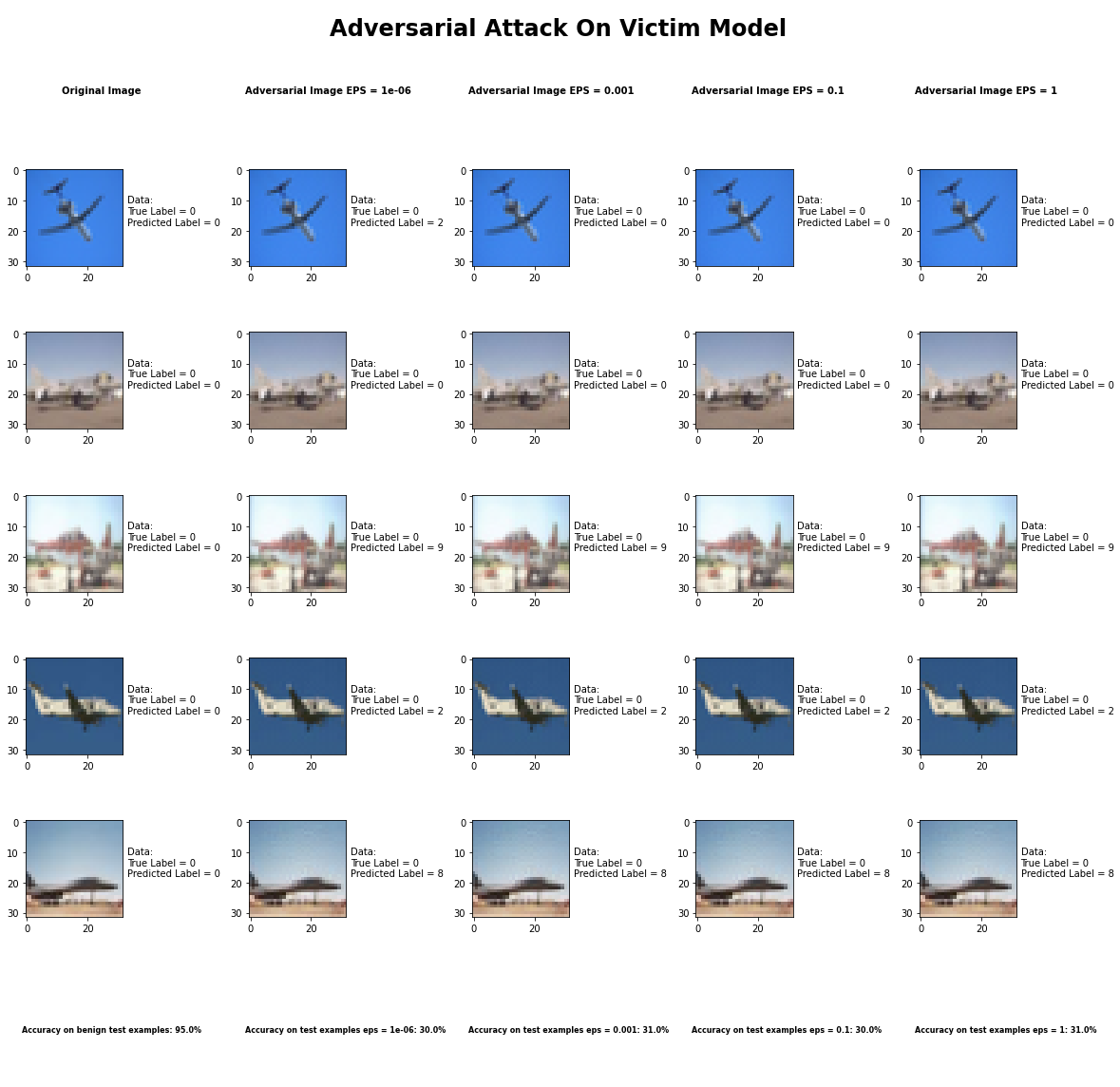
Universal Perturbation Method

|  |  |
| --- | --- |
| Average Accuracy Over 10 Runs | |
| Benign Samples | 94.52% |
| Epsilon = 0.05 Samples | 48.4% |
| Epsilon = 0.10 Samples | 39.9% |
| Epsilon = 0.5 Samples | 35.3% |
| Epsilon = 0.95 Samples | 35.3% |

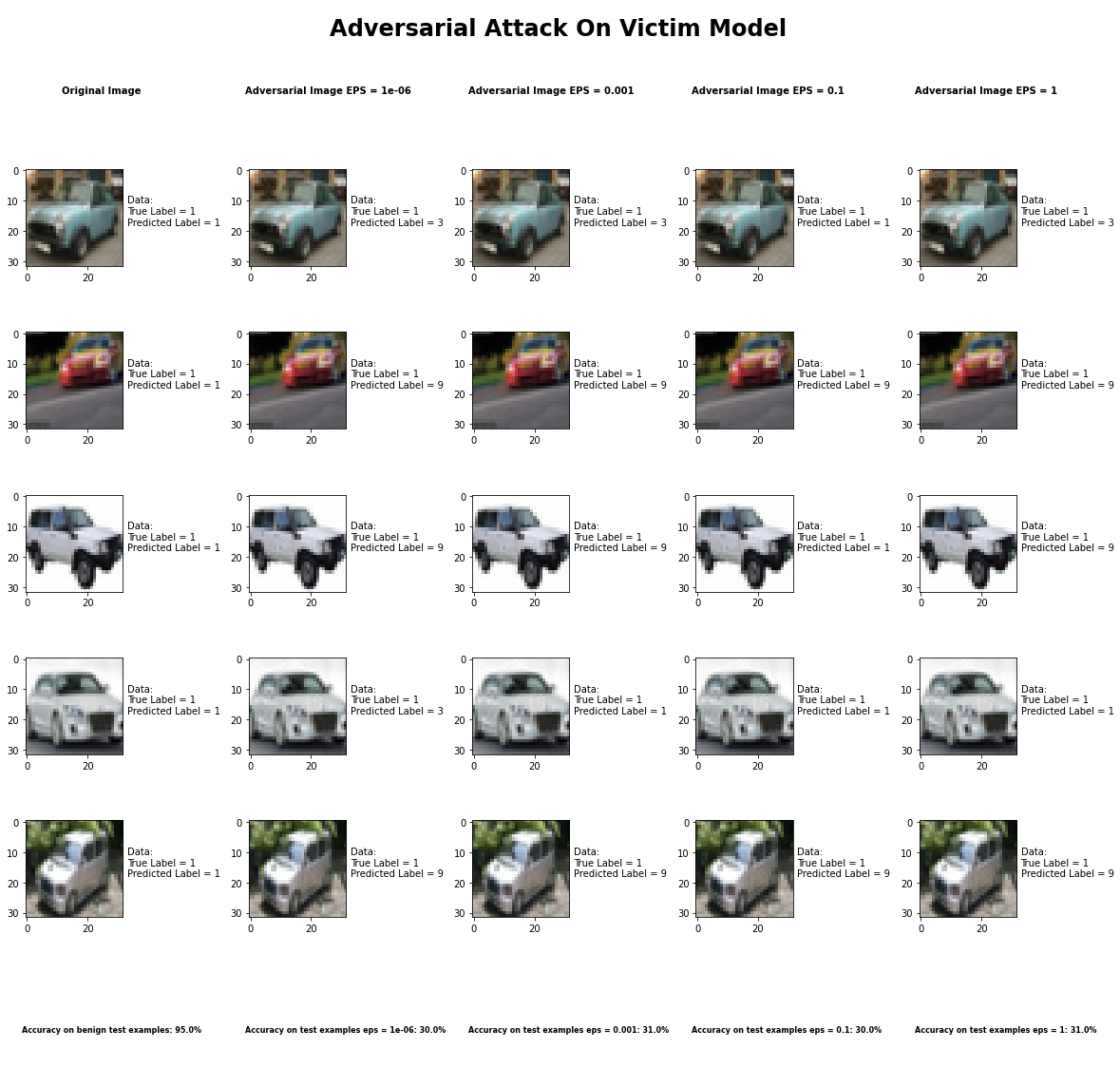
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Average Accuracy per Class Over 10 Runs | | | | | | | | | | |
|  | Airplane | Car | Bird | Cat | Deer | Dog | Frog | Horse | Ship | Truck |
| Benign Samples | 98% | 96% | 95% | 86% | 96% | 84% | 96% | 93% | 99% | 97% |
| Epsilon = 0.05 Samples | 32% | 60% | 68% | 42% | 43% | 31% | 60% | 33% | 58% | 57% |
| Epsilon = 0.10 Samples | 30% | 40% | 70% | 34% | 28% | 23% | 68% | 24% | 45% | 37% |
| Epsilon = 0.5 Samples | 33% | 40% | 34% | 37% | 20% | 23% | 26% | 24% | 45% | 71% |
| Epsilon = 0.95 Samples | 30% | 35% | 34% | 32% | 20% | 23% | 26% | 23% | 45% | 87% |

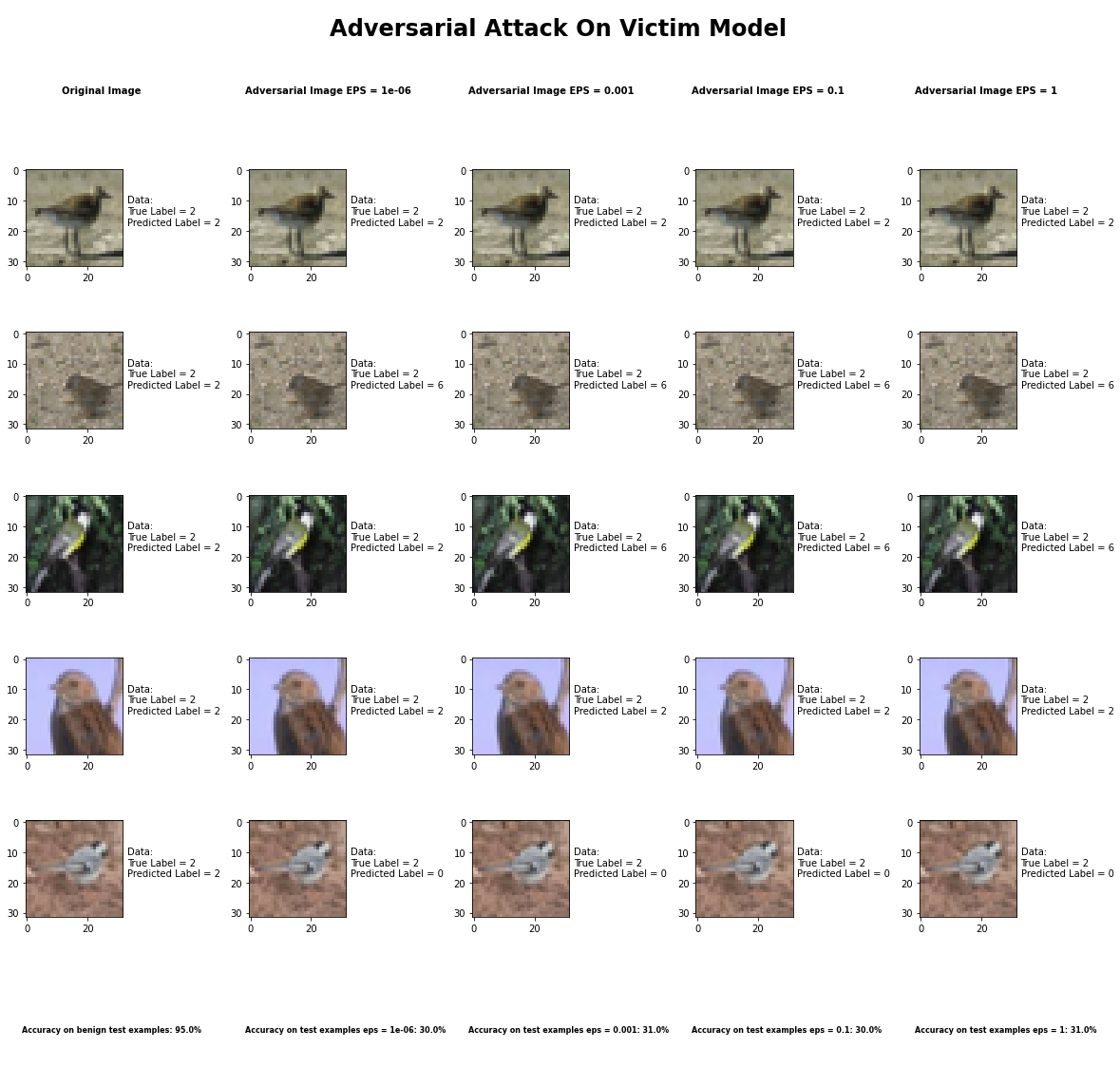
DeepFool Method

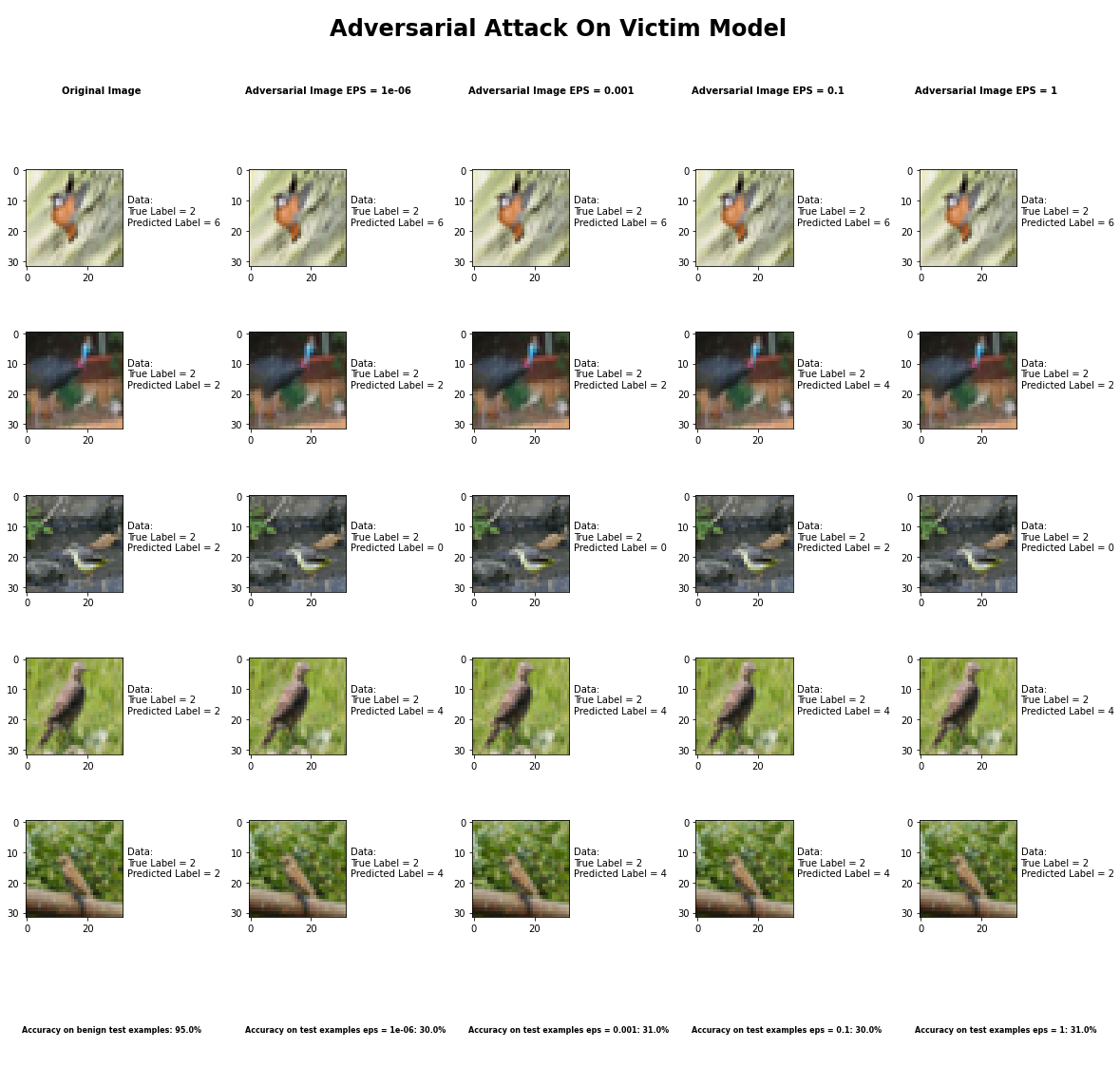




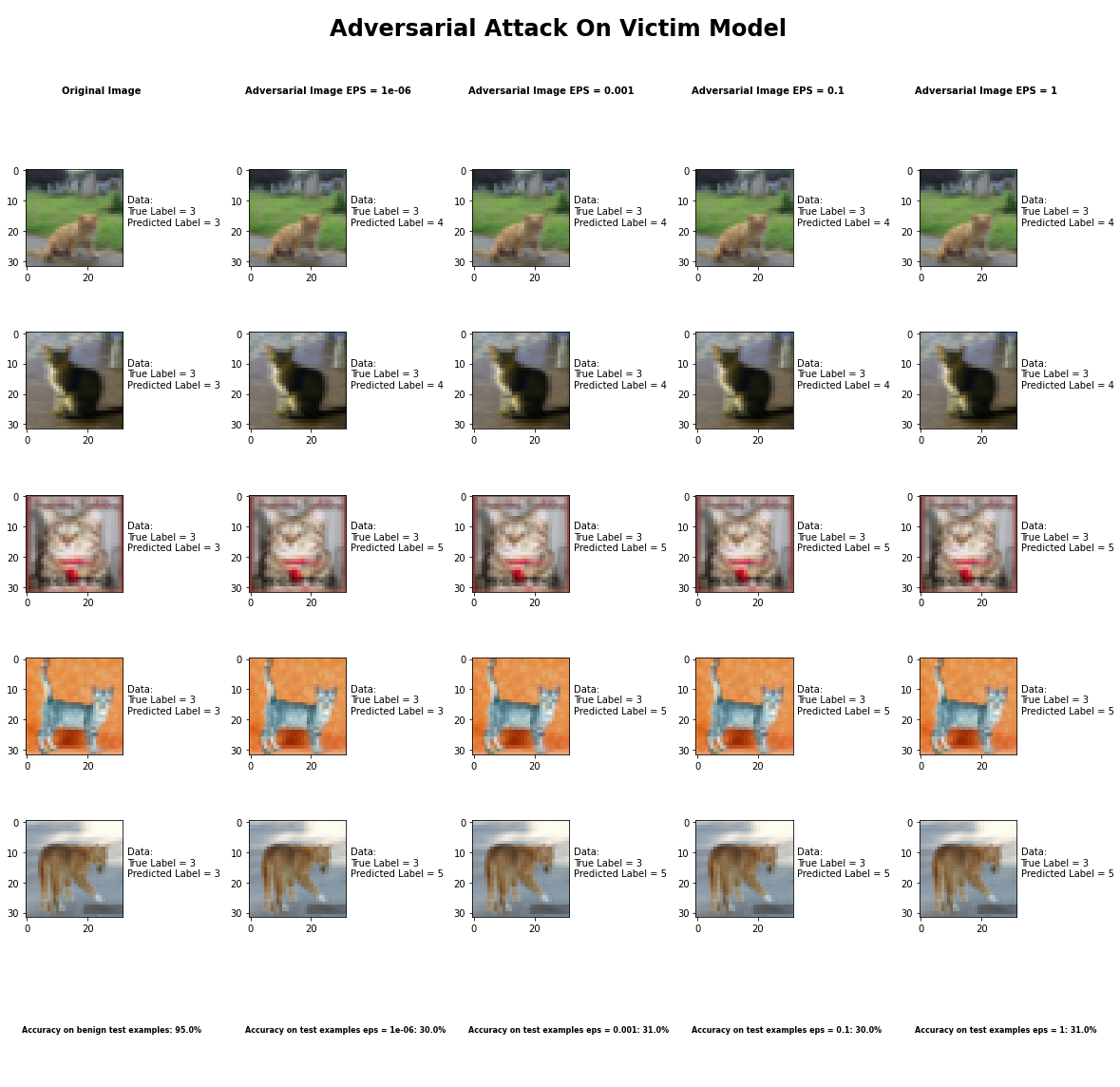




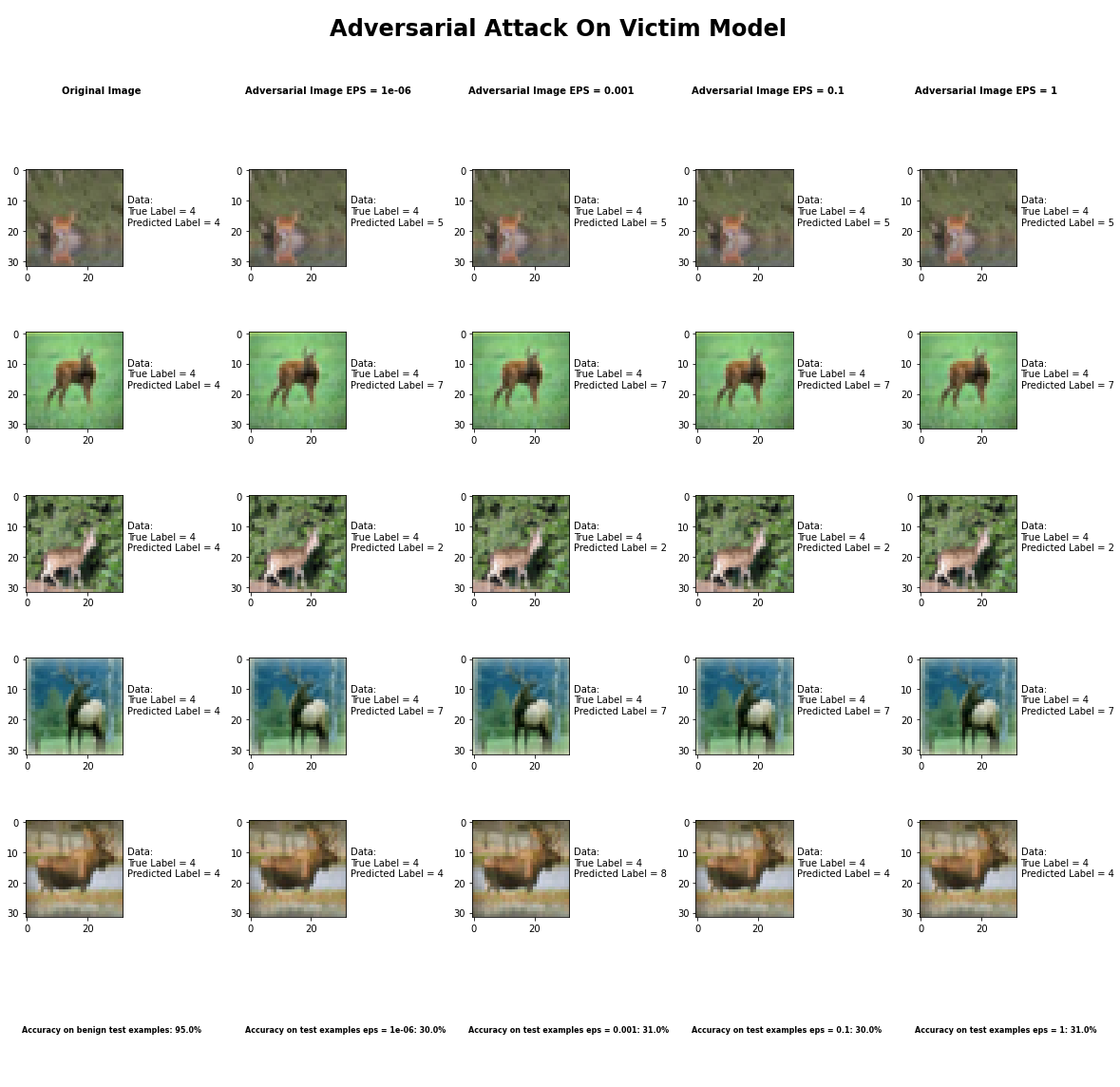








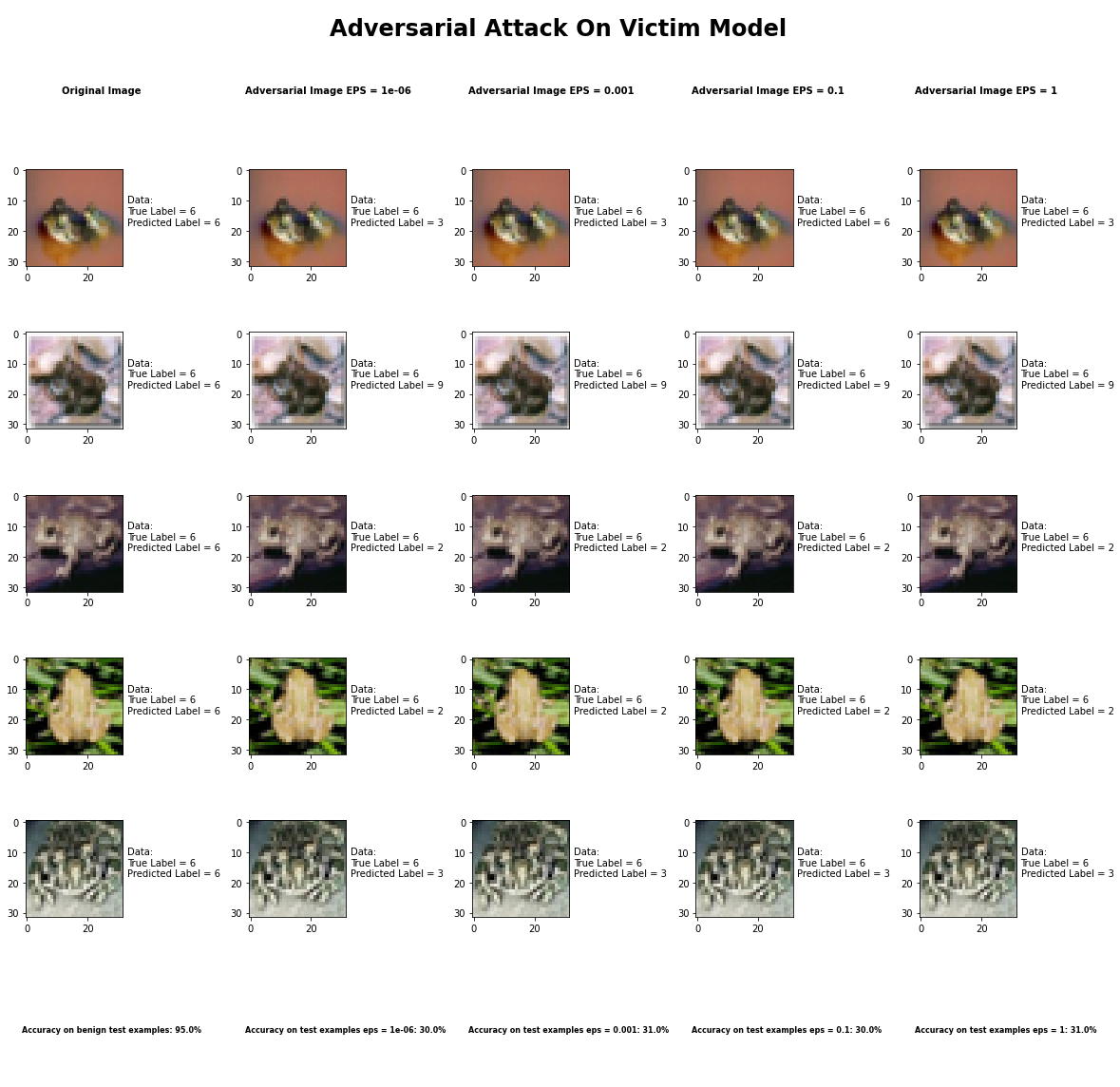




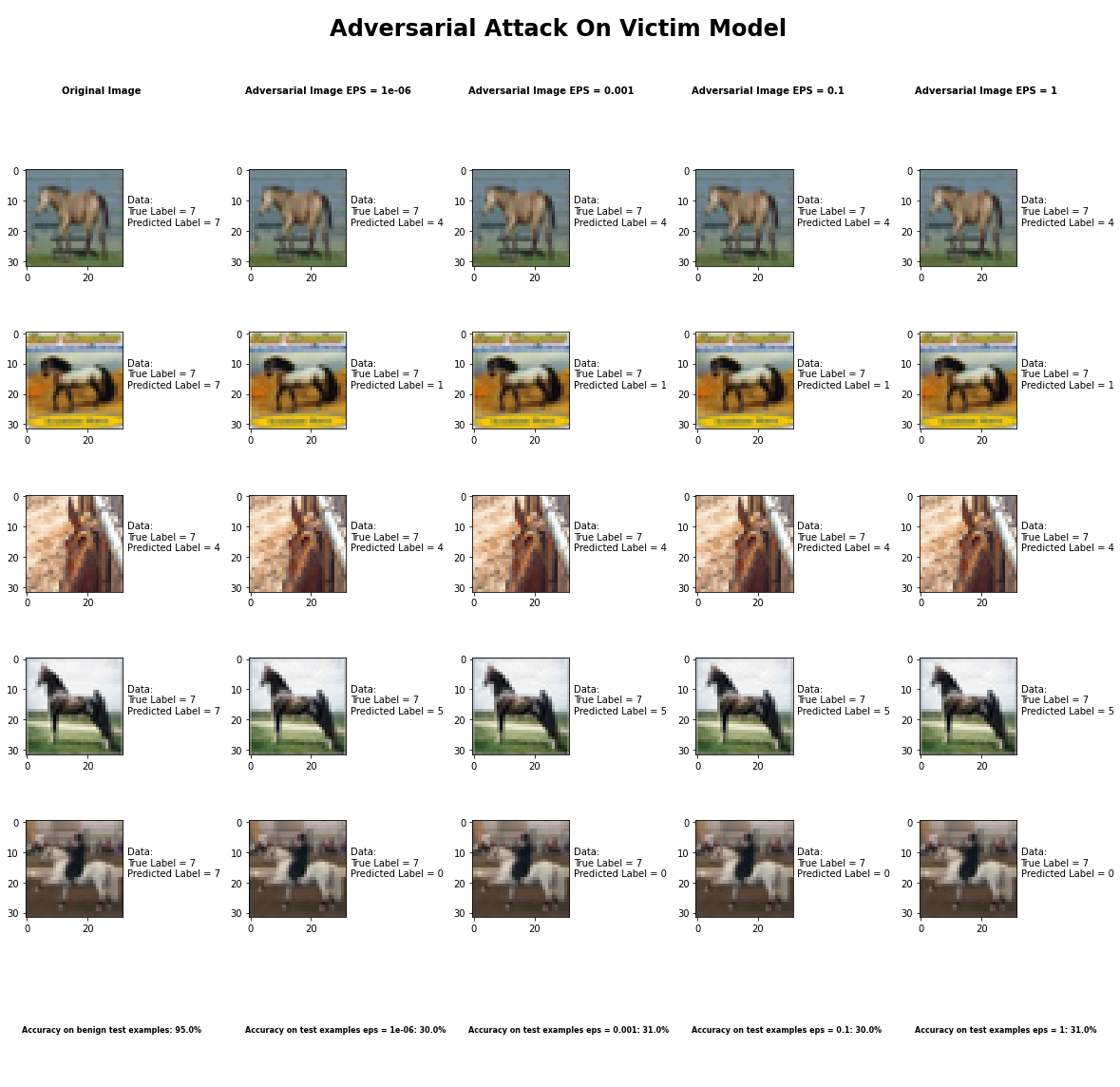


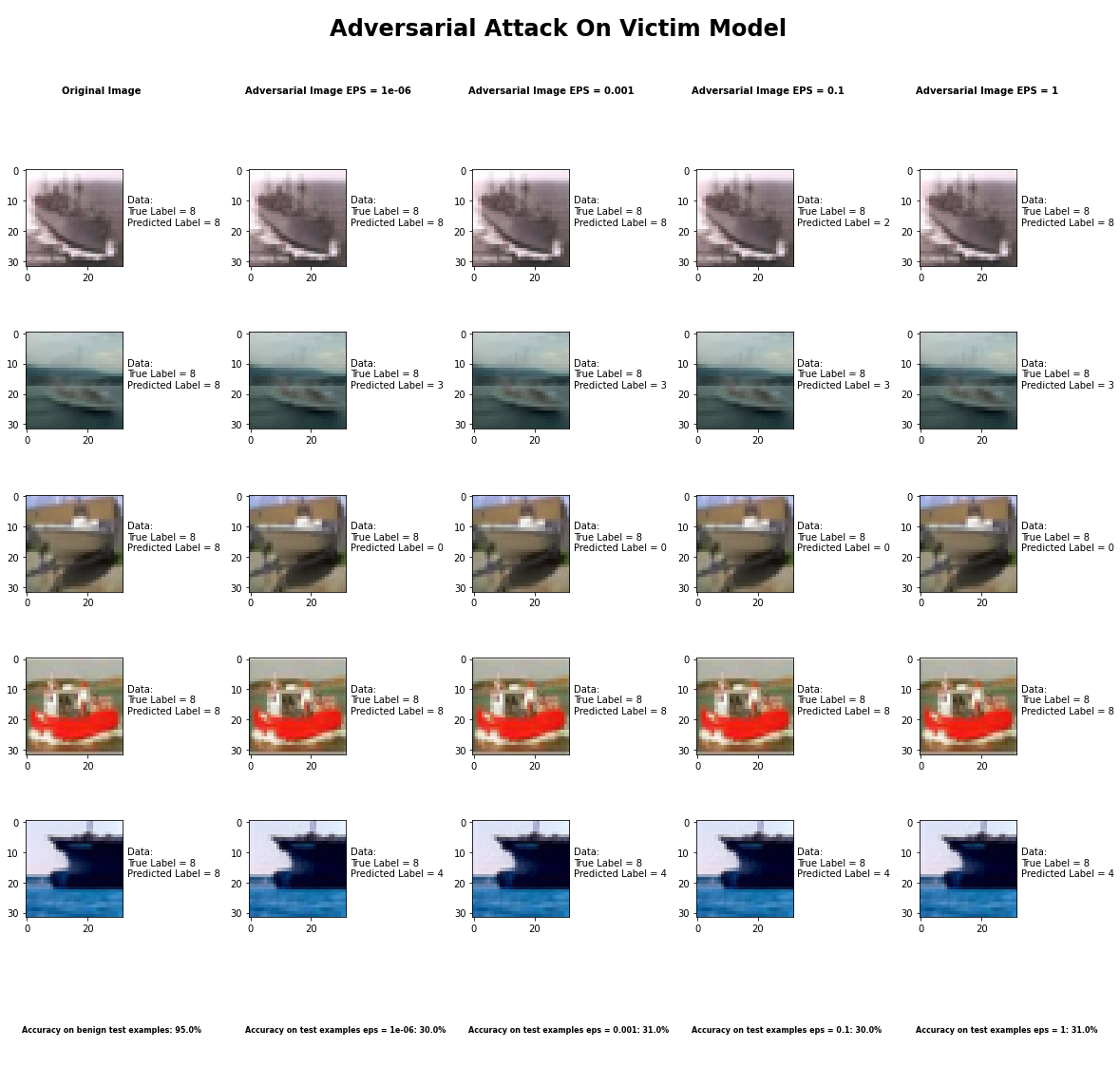


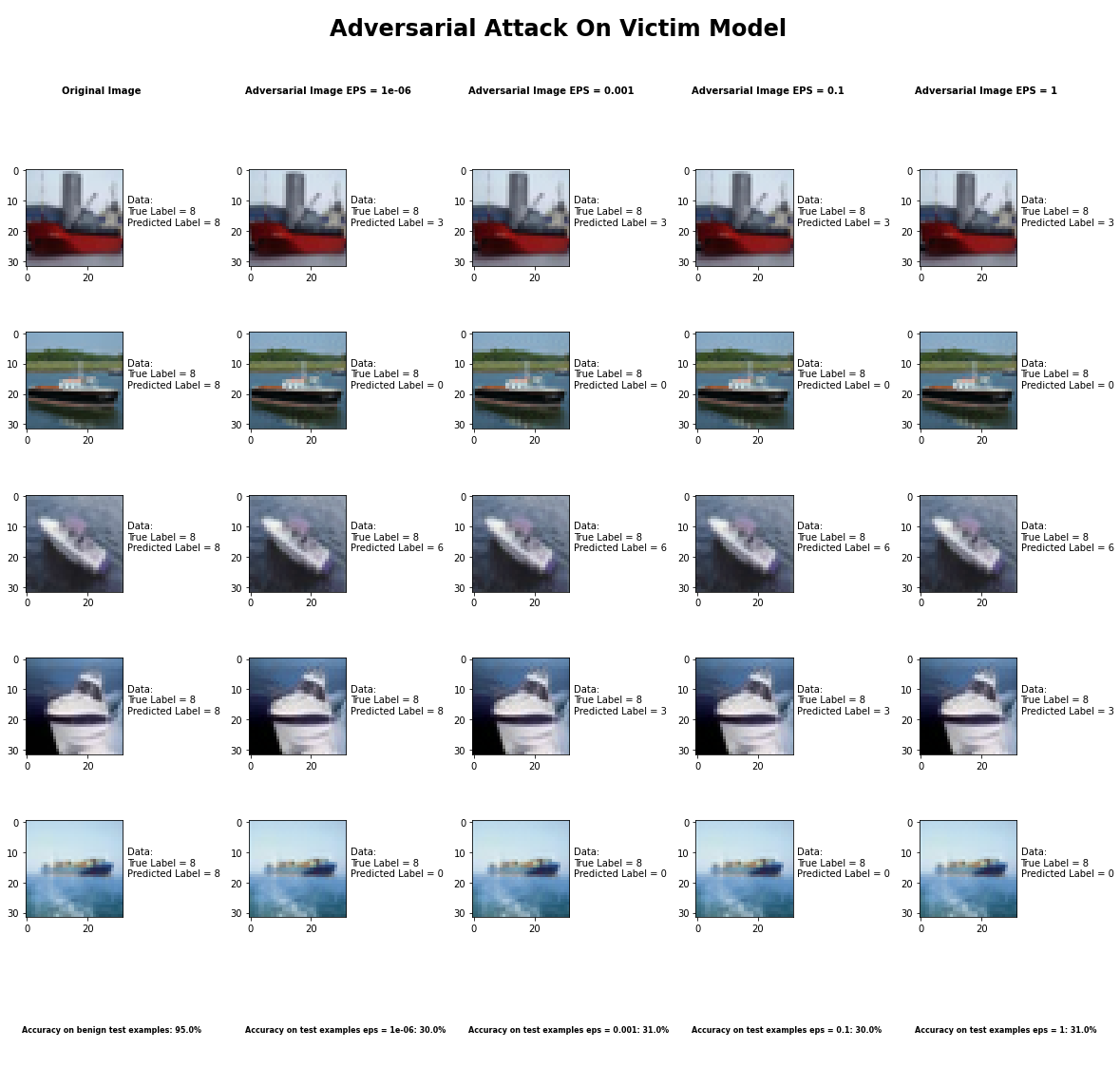


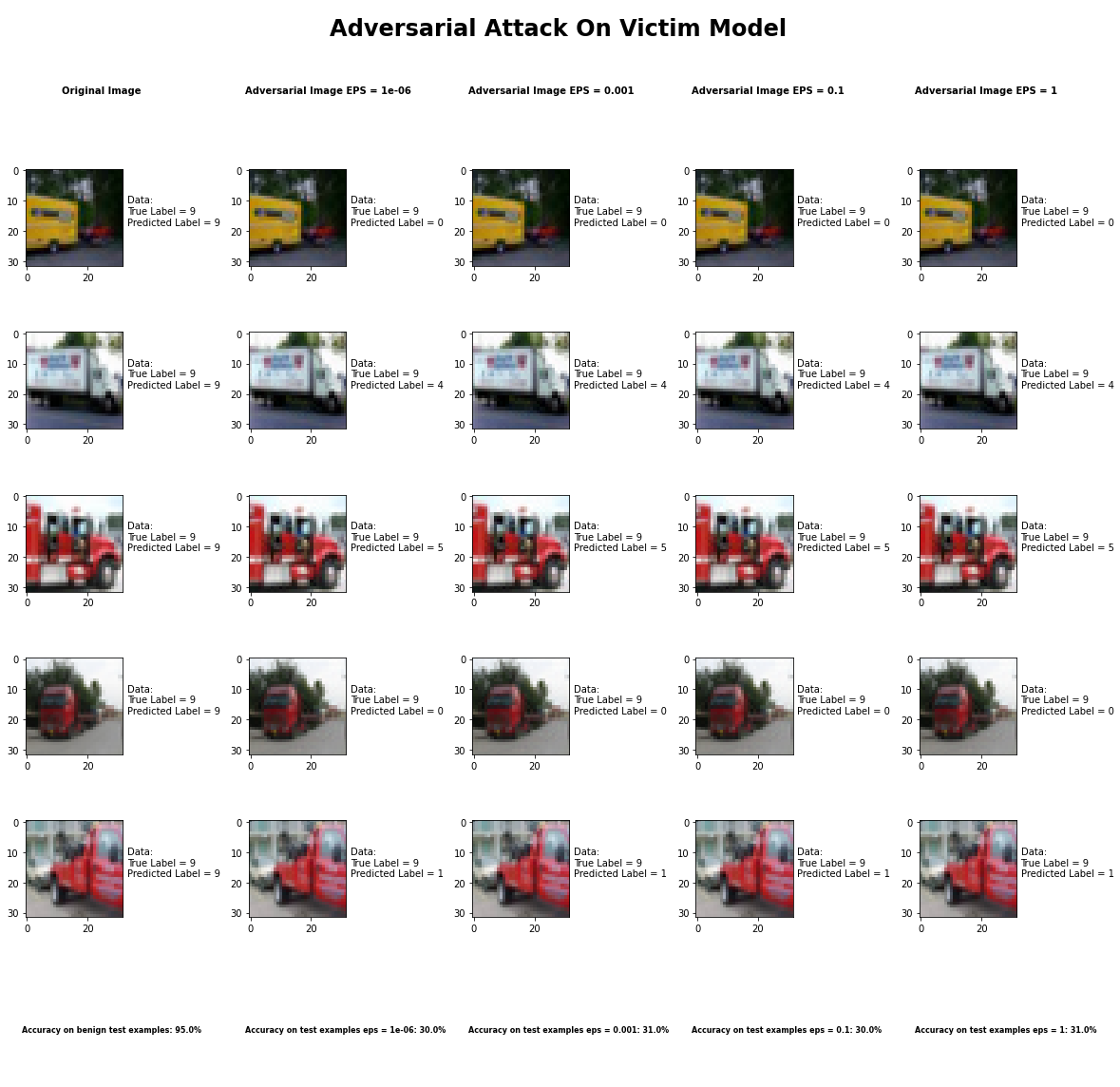


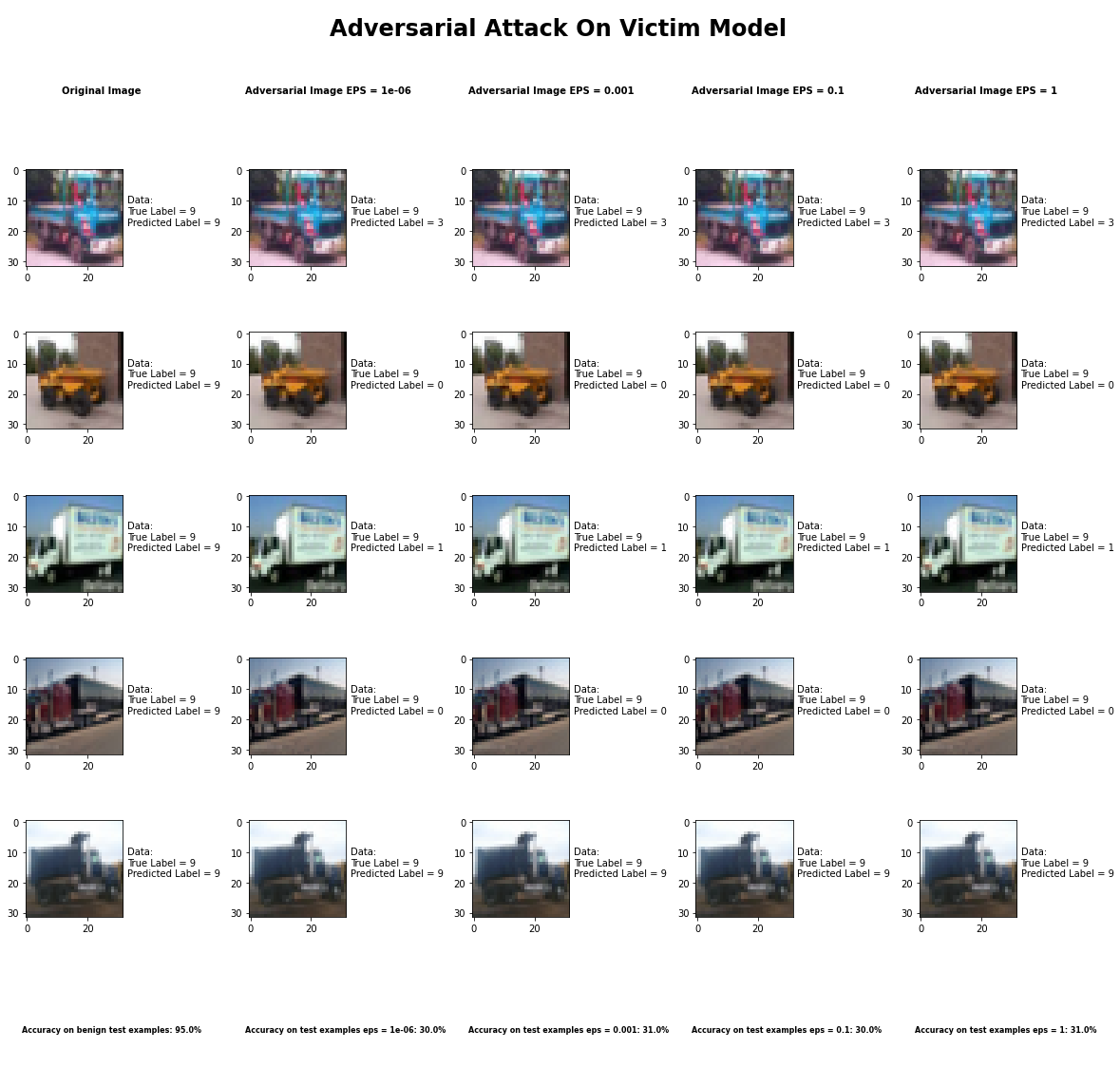












Accuracy on benign test examples: 94.52000000000001%

Average Accuracy on adversarial test examples with eps = 0.05: 30.2%

Average Accuracy on adversarial test examples with eps = 0.1: 31.3%

Average Accuracy on adversarial test examples with eps = 0.5: 30.1%

Average Accuracy on adversarial test examples with eps = 0.95: 31.0%

Classifier with benign example has Airplane recognition average accuracy of = 95%

DeepFool Method with eps = 1e-6 has Airplane recognition average accuracy of = 42%

DeepFool Method with eps = 1e-3 has Airplane recognition average accuracy of = 46%

DeepFool Method with eps = 1e-1 has Airplane recognition average accuracy of = 46%

DeepFool Method with eps = 1 has Airplane recognition average accuracy of = 46%

Classifier with benign example has Automobile recognition average accuracy of = 92%

DeepFool Method with eps = 1e-6 has Automobile recognition average accuracy of = 45%

DeepFool Method with eps = 1e-3 has Automobile recognition average accuracy of = 51%

DeepFool Method with eps = 1e-1 has Automobile recognition average accuracy of = 48%

DeepFool Method with eps = 1 has Automobile recognition average accuracy of = 41%

Classifier with benign example has Bird recognition average accuracy of = 96%

DeepFool Method with eps = 1e-6 has Bird recognition average accuracy of = 27%

DeepFool Method with eps = 1e-3 has Bird recognition average accuracy of = 31%

DeepFool Method with eps = 1e-1 has Bird recognition average accuracy of = 24%

DeepFool Method with eps = 1 has Bird recognition average accuracy of = 28%

Classifier with benign example has Cat recognition average accuracy of = 86%

DeepFool Method with eps = 1e-6 has Cat recognition average accuracy of = 29%

DeepFool Method with eps = 1e-3 has Cat recognition average accuracy of = 26%

DeepFool Method with eps = 1e-1 has Cat recognition average accuracy of = 22%

DeepFool Method with eps = 1 has Cat recognition average accuracy of = 24%

Classifier with benign example has Deer recognition average accuracy of = 95%

DeepFool Method with eps = 1e-6 has Deer recognition average accuracy of = 37%

DeepFool Method with eps = 1e-3 has Deer recognition average accuracy of = 33%

DeepFool Method with eps = 1e-1 has Deer recognition average accuracy of = 38%

DeepFool Method with eps = 1 has Deer recognition average accuracy of = 35%

Classifier with benign example has Dog recognition average accuracy of = 88%

DeepFool Method with eps = 1e-6 has Dog recognition average accuracy of = 20%

DeepFool Method with eps = 1e-3 has Dog recognition average accuracy of = 22%

DeepFool Method with eps = 1e-1 has Dog recognition average accuracy of = 23%

DeepFool Method with eps = 1 has Dog recognition average accuracy of = 22%

Classifier with benign example has Frog recognition average accuracy of = 99%

DeepFool Method with eps = 1e-6 has Frog recognition average accuracy of = 26%

DeepFool Method with eps = 1e-3 has Frog recognition average accuracy of = 24%

DeepFool Method with eps = 1e-1 has Frog recognition average accuracy of = 29%

DeepFool Method with eps = 1 has Frog recognition average accuracy of = 25%

Classifier with benign example has Horse recognition average accuracy of = 93%

DeepFool Method with eps = 1e-6 has Horse recognition average accuracy of = 21%

DeepFool Method with eps = 1e-3 has Horse recognition average accuracy of = 25%

DeepFool Method with eps = 1e-1 has Horse recognition average accuracy of = 24%

DeepFool Method with eps = 1 has Horse recognition average accuracy of = 26%

Classifier with benign example has Ship recognition average accuracy of = 94%

DeepFool Method with eps = 1e-6 has Ship recognition average accuracy of = 32%

DeepFool Method with eps = 1e-3 has Ship recognition average accuracy of = 30%

DeepFool Method with eps = 1e-1 has Ship recognition average accuracy of = 28%

DeepFool Method with eps = 1 has Ship recognition average accuracy of = 32%

Classifier with benign example has Truck recognition average accuracy of = 98%

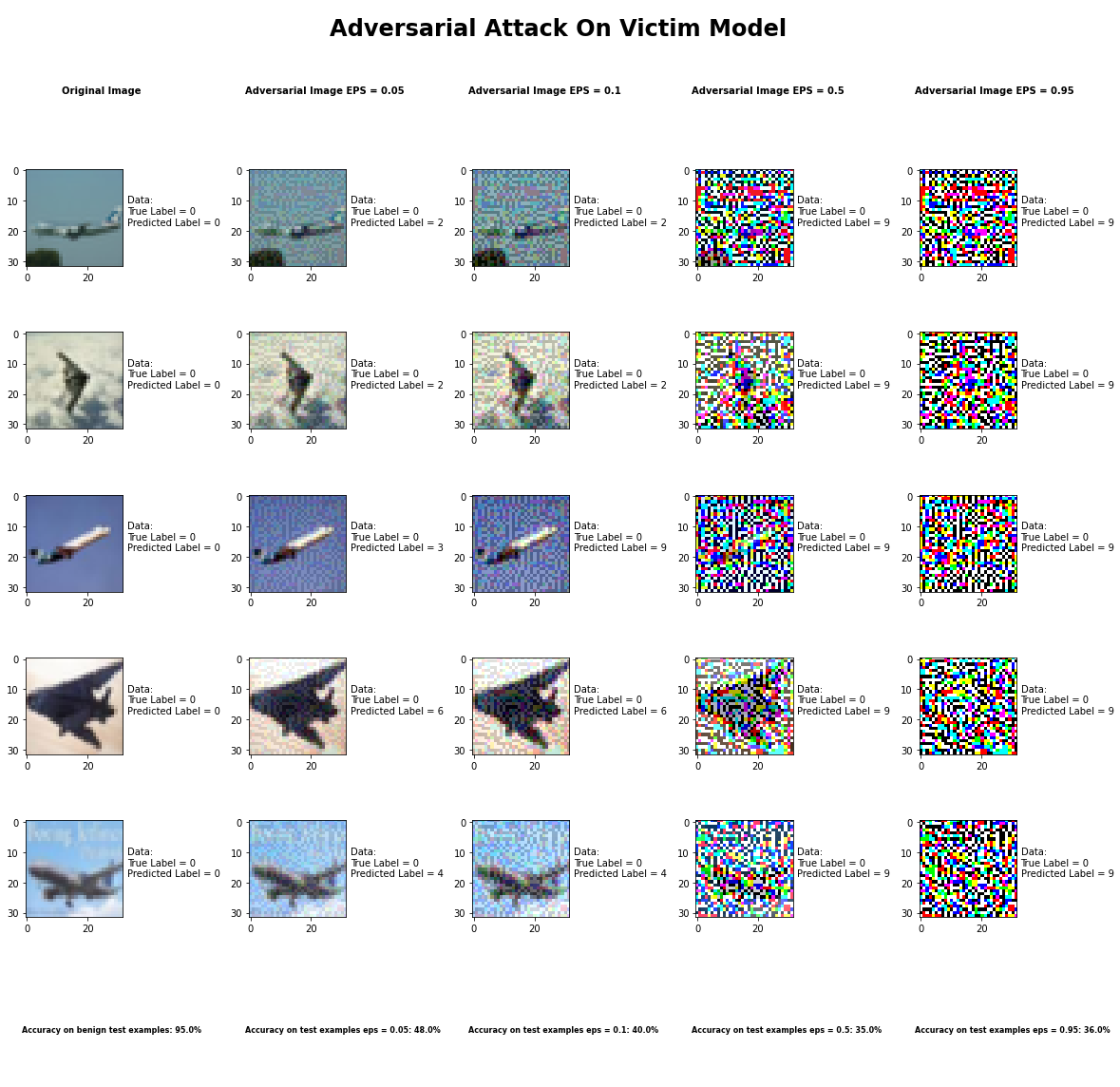
DeepFool Method with eps = 1e-6 has Truck recognition average accuracy of = 26%

DeepFool Method with eps = 1e-3 has Truck recognition average accuracy of = 24%

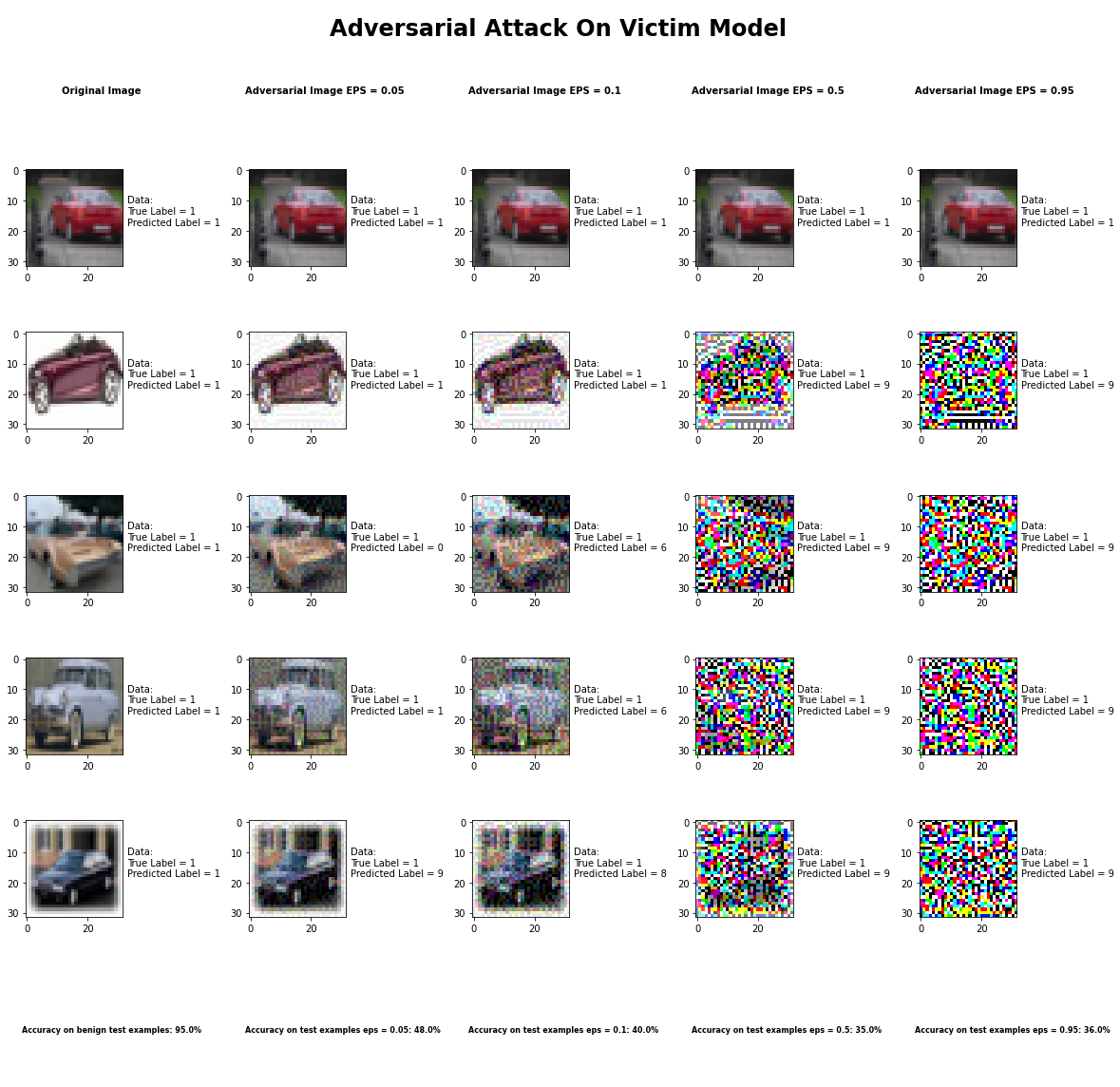
DeepFool Method with eps = 1e-1 has Truck recognition average accuracy of = 21%

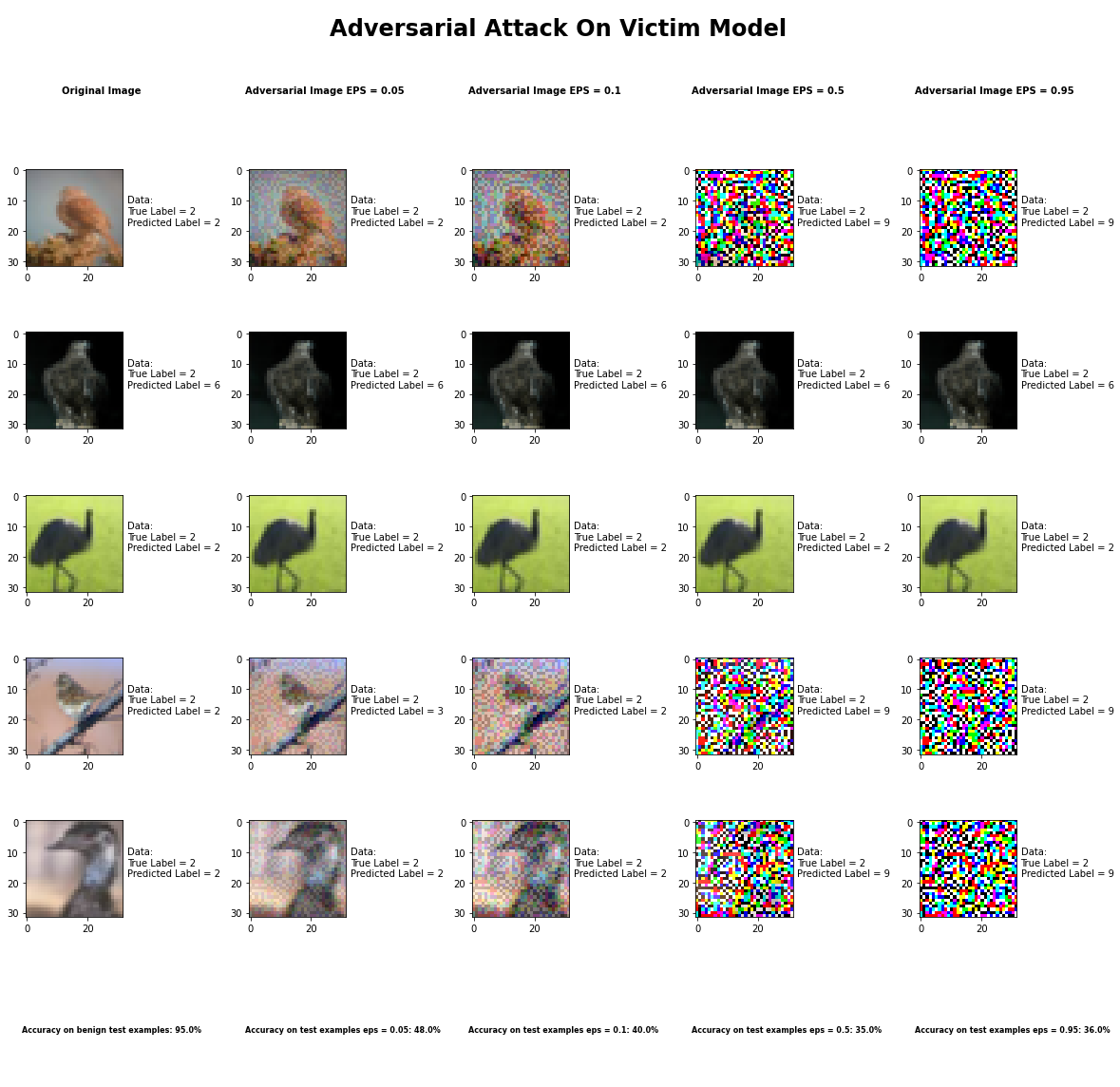
DeepFool Method with eps = 1 has Truck recognition average accuracy of = 18%

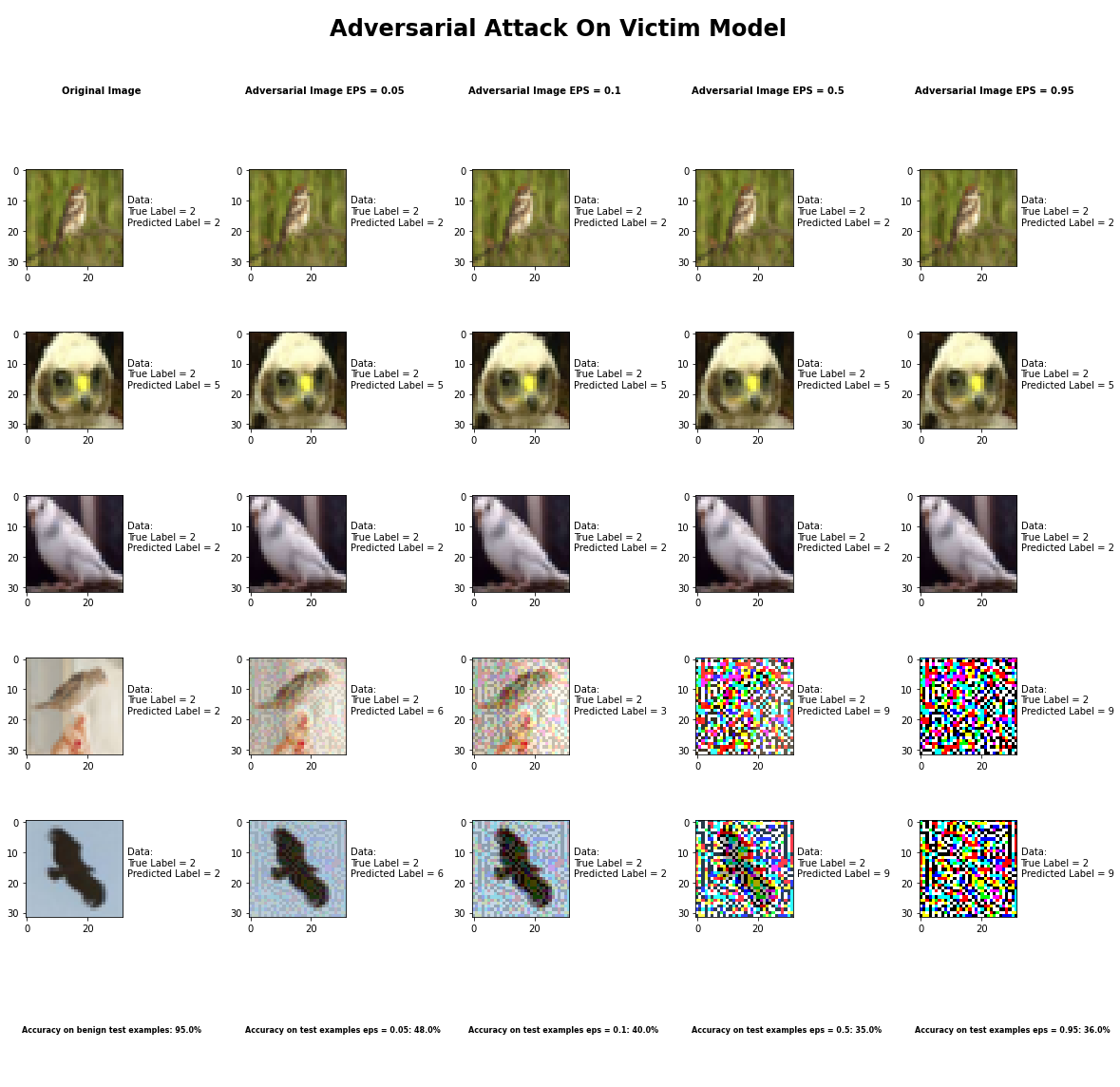
Fast Gradient Sign Method

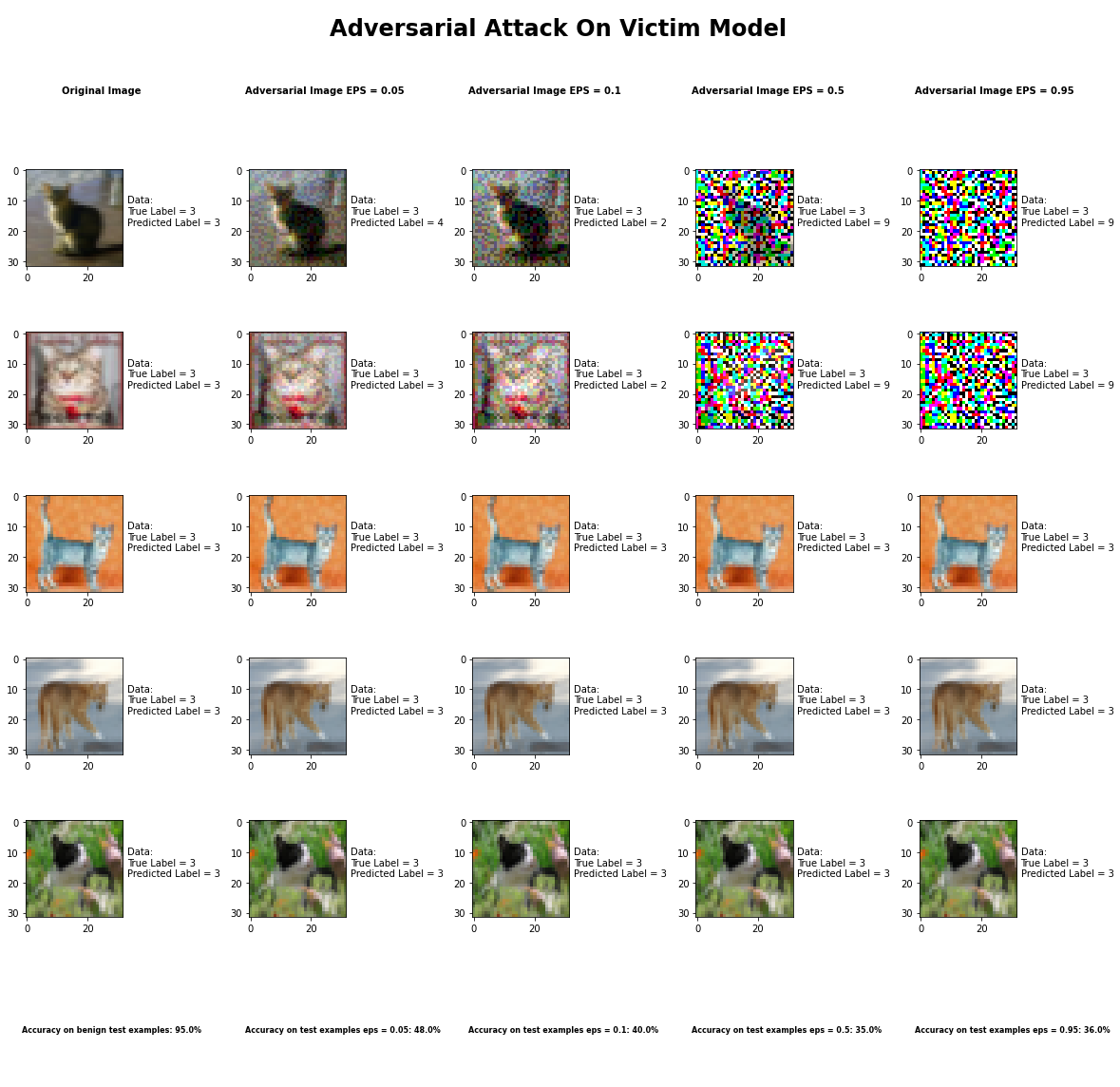






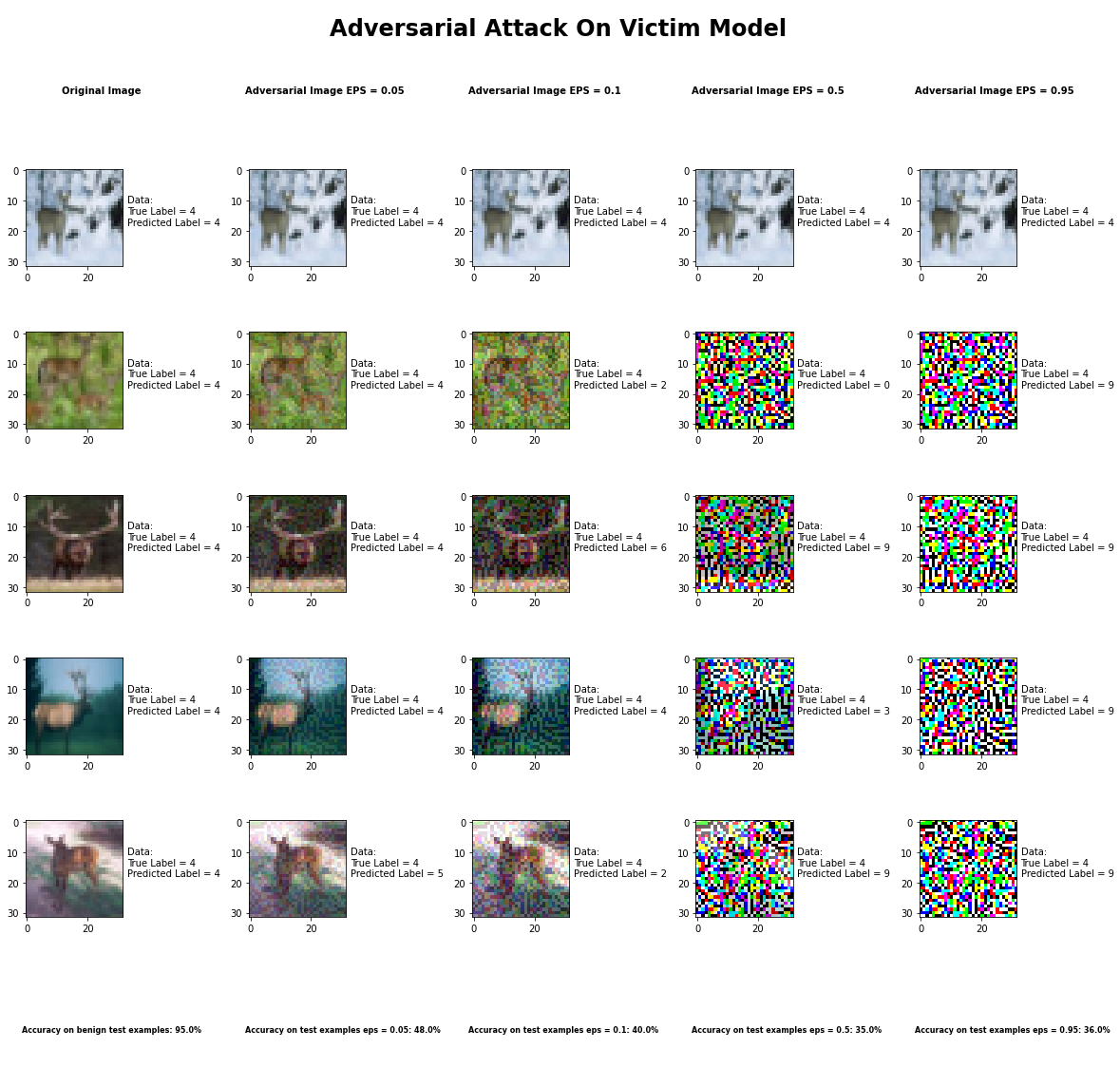


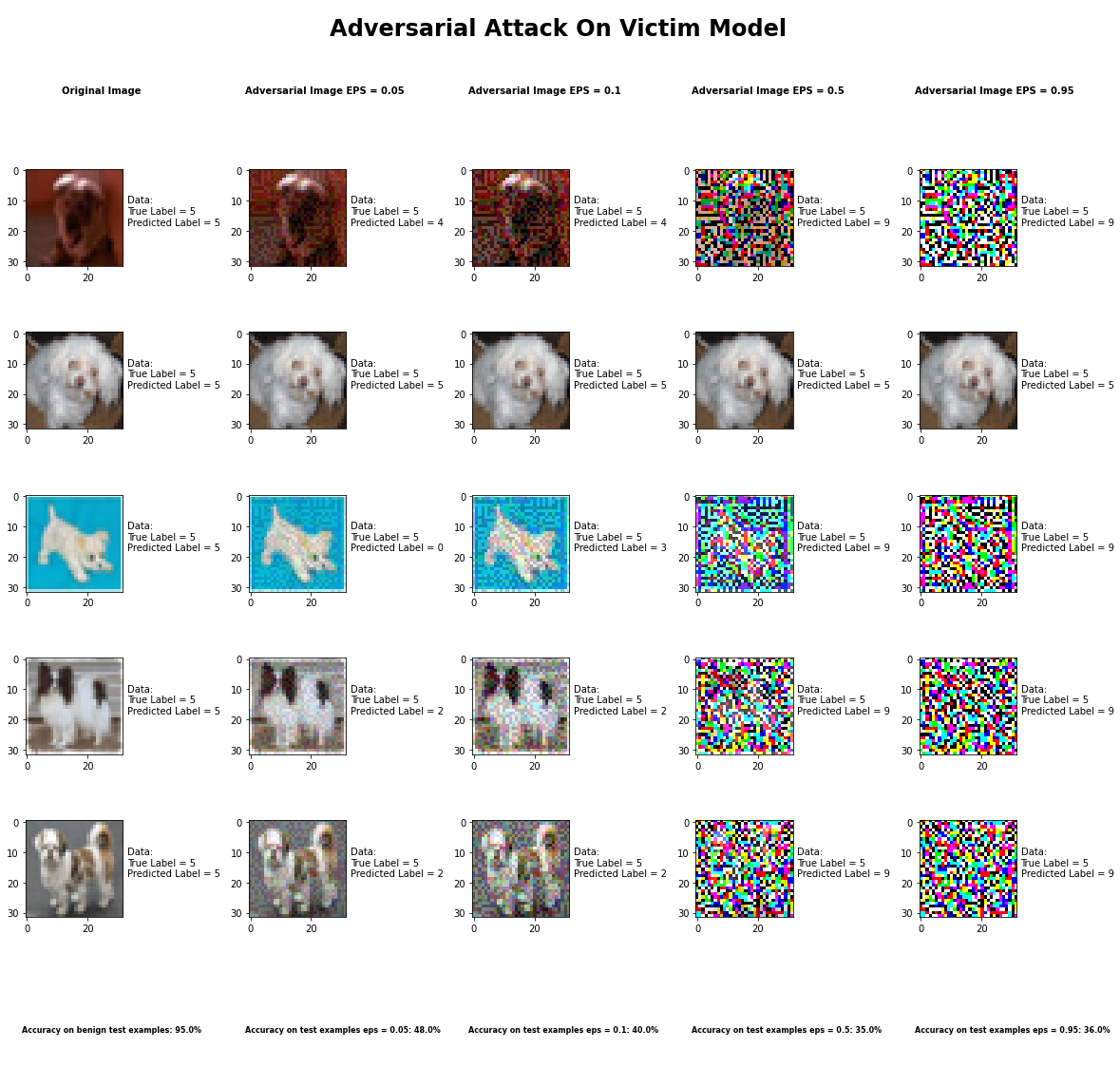






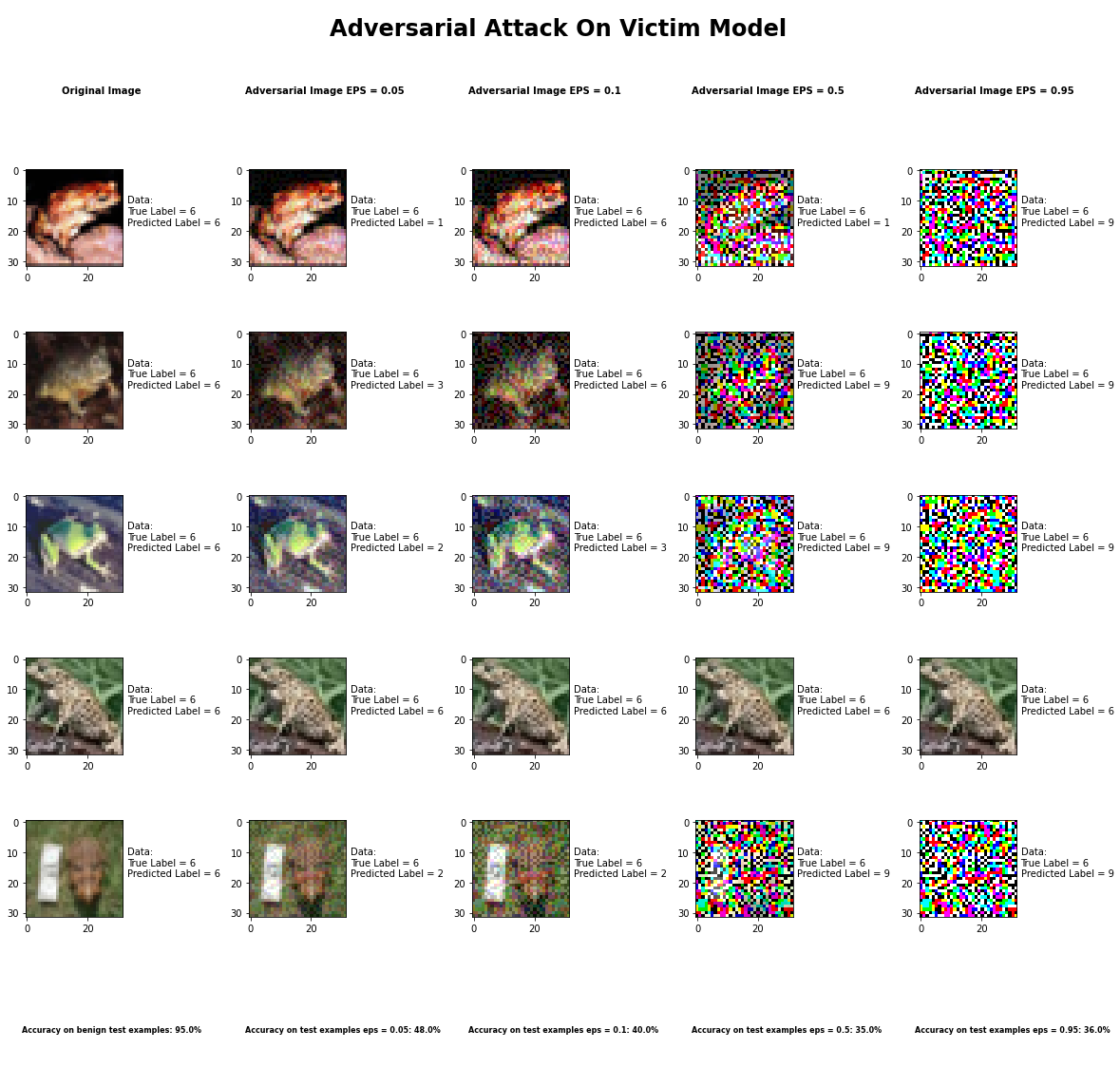






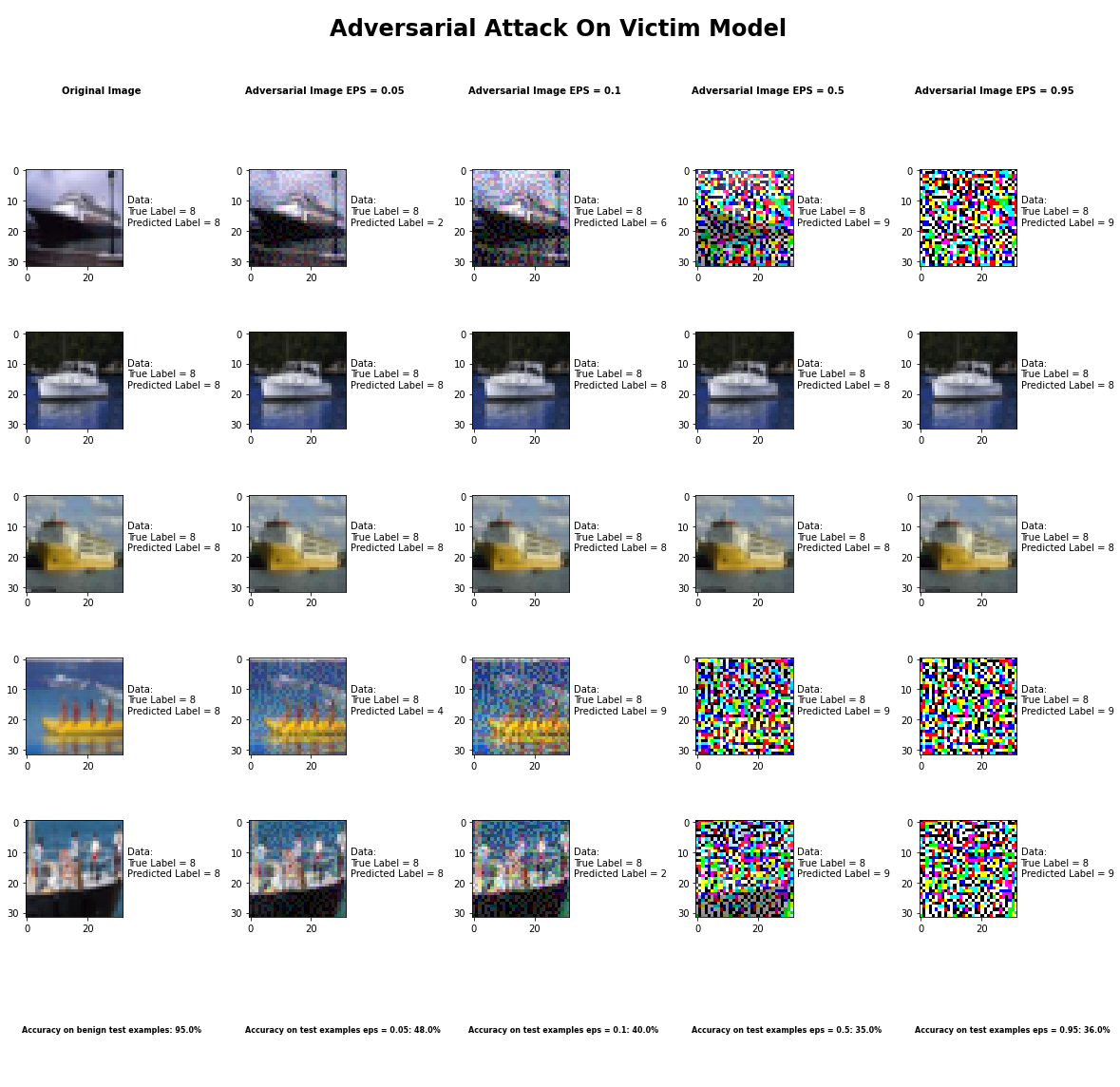


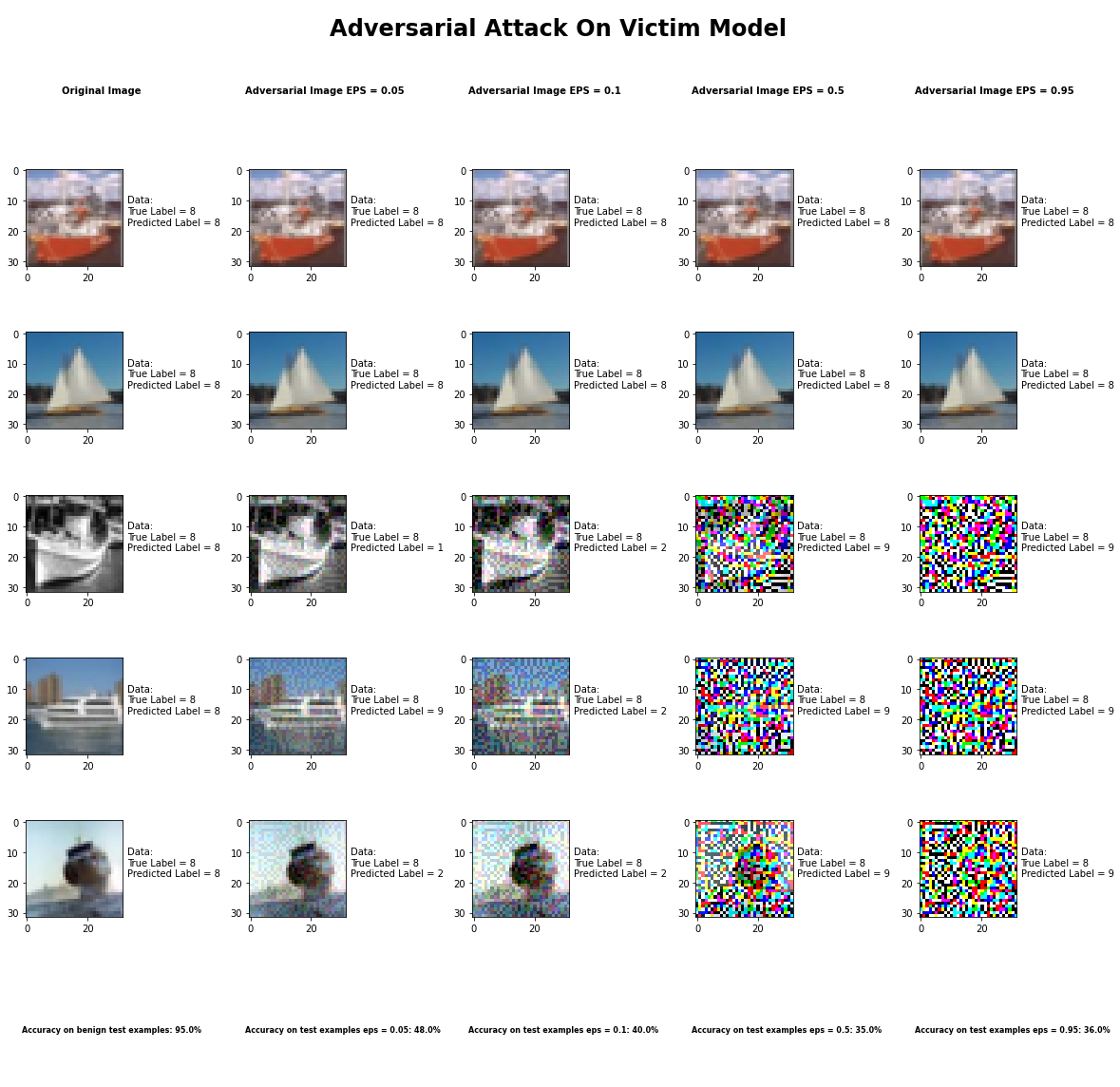
















Accuracy on benign test examples: 94.52000000000001%

Average Accuracy on adversarial test examples with eps = 0.05: 48.4%

Average Accuracy on adversarial test examples with eps = 0.1: 39.9%

Average Accuracy on adversarial test examples with eps = 0.5: 35.3%

Average Accuracy on adversarial test examples with eps = 0.95: 35.5%

Classifier with benign example has Airplane recognition average accuracy of = 98%

Fast Gradient Method with eps = 0.05 has Airplane recognition average accuracy of = 32%

Fast Gradient Method with eps = 0.10 has Airplane recognition average accuracy of = 30%

Fast Gradient Method with eps = 0.50 has Airplane recognition average accuracy of = 33%

Fast Gradient Method with eps = 0.95 has Airplane recognition average accuracy of = 30%

Classifier with benign example has Automobile recognition average accuracy of = 96%

Fast Gradient Method with eps = 0.05 has Automobile recognition average accuracy of = 60%

Fast Gradient Method with eps = 0.10 has Automobile recognition average accuracy of = 40%

Fast Gradient Method with eps = 0.50 has Automobile recognition average accuracy of = 40%

Fast Gradient Method with eps = 0.95 has Automobile recognition average accuracy of = 35%

Classifier with benign example has Bird recognition average accuracy of = 95%

Fast Gradient Method with eps = 0.05 has Bird recognition average accuracy of = 68%

Fast Gradient Method with eps = 0.10 has Bird recognition average accuracy of = 70%

Fast Gradient Method with eps = 0.50 has Bird recognition average accuracy of = 34%

Fast Gradient Method with eps = 0.95 has Bird recognition average accuracy of = 34%

Classifier with benign example has Cat recognition average accuracy of = 86%

Fast Gradient Method with eps = 0.05 has Cat recognition average accuracy of = 42%

Fast Gradient Method with eps = 0.10 has Cat recognition average accuracy of = 34%

Fast Gradient Method with eps = 0.50 has Cat recognition average accuracy of = 37%

Fast Gradient Method with eps = 0.95 has Cat recognition average accuracy of = 32%

Classifier with benign example has Deer recognition average accuracy of = 96%

Fast Gradient Method with eps = 0.05 has Deer recognition average accuracy of = 43%

Fast Gradient Method with eps = 0.10 has Deer recognition average accuracy of = 28%

Fast Gradient Method with eps = 0.50 has Deer recognition average accuracy of = 20%

Fast Gradient Method with eps = 0.95 has Deer recognition average accuracy of = 20%

Classifier with benign example has Dog recognition average accuracy of = 84%

Fast Gradient Method with eps = 0.05 has Dog recognition average accuracy of = 31%

Fast Gradient Method with eps = 0.10 has Dog recognition average accuracy of = 23%

Fast Gradient Method with eps = 0.50 has Dog recognition average accuracy of = 23%

Fast Gradient Method with eps = 0.95 has Dog recognition average accuracy of = 23%

Classifier with benign example has Frog recognition average accuracy of = 96%

Fast Gradient Method with eps = 0.05 has Frog recognition average accuracy of = 60%

Fast Gradient Method with eps = 0.10 has Frog recognition average accuracy of = 68%

Fast Gradient Method with eps = 0.50 has Frog recognition average accuracy of = 26%

Fast Gradient Method with eps = 0.95 has Frog recognition average accuracy of = 26%

Classifier with benign example has Horse recognition average accuracy of = 93%

Fast Gradient Method with eps = 0.05 has Horse recognition average accuracy of = 33%

Fast Gradient Method with eps = 0.10 has Horse recognition average accuracy of = 24%

Fast Gradient Method with eps = 0.50 has Horse recognition average accuracy of = 24%

Fast Gradient Method with eps = 0.95 has Horse recognition average accuracy of = 23%

Classifier with benign example has Ship recognition average accuracy of = 99%

Fast Gradient Method with eps = 0.05 has Ship recognition average accuracy of = 58%

Fast Gradient Method with eps = 0.10 has Ship recognition average accuracy of = 45%

Fast Gradient Method with eps = 0.50 has Ship recognition average accuracy of = 45%

Fast Gradient Method with eps = 0.95 has Ship recognition average accuracy of = 45%

Classifier with benign example has Truck recognition average accuracy of = 97%

Fast Gradient Method with eps = 0.05 has Truck recognition average accuracy of = 57%

Fast Gradient Method with eps = 0.10 has Truck recognition average accuracy of = 37%

Fast Gradient Method with eps = 0.50 has Truck recognition average accuracy of = 71%

Fast Gradient Method with eps = 0.95 has Truck recognition average accuracy of = 87%