HA6 Report

Name : QIU Yaowen
ID : 20784389

Pre-Trained Model:

MobileNetv2

Predictive Score:

ID	Probability	Predicted Label
1	0.3622	Rapeseed
2	0.9913	Peacock
3	0.9297	Yurt
4	0.9724	Hourglass
5	0.9453	Water Tower
6	0.9862	Zebra
7	0.9862	School Bus
8	0.9978	Pillow
9	0.7541	Fireboat
10	0.9732	Carousel

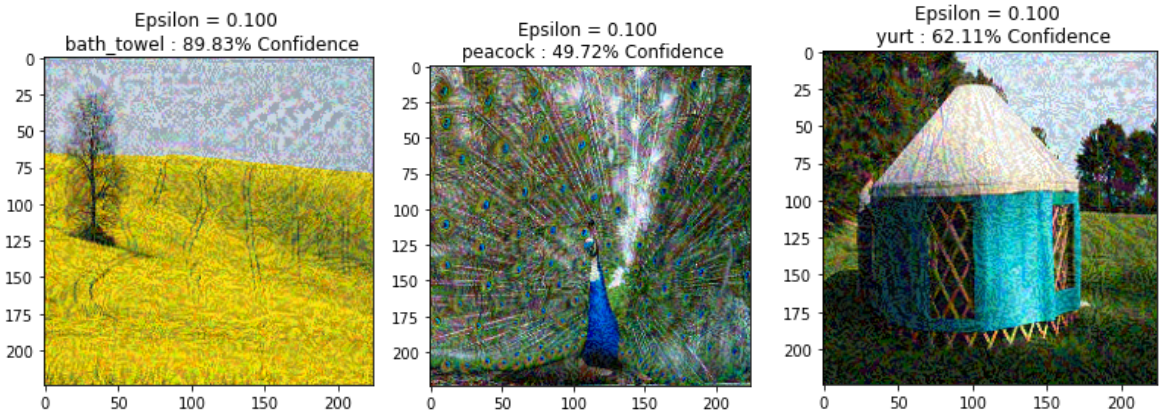
Adversarial attack Method

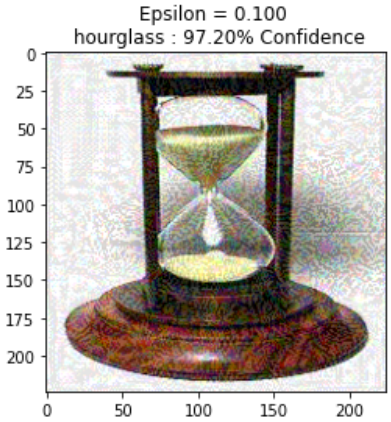
The fast gradient sign method is used, since it is fast and popular.

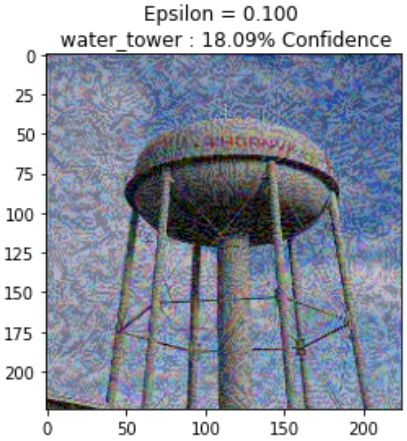
Parameter Setting:

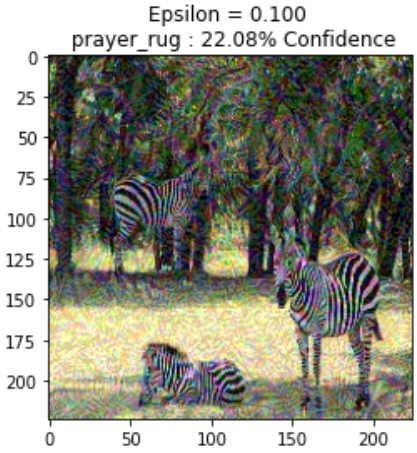
Epsilon is set to be 0.1 in this experiment, which still makes the new images quite similar to the original one.

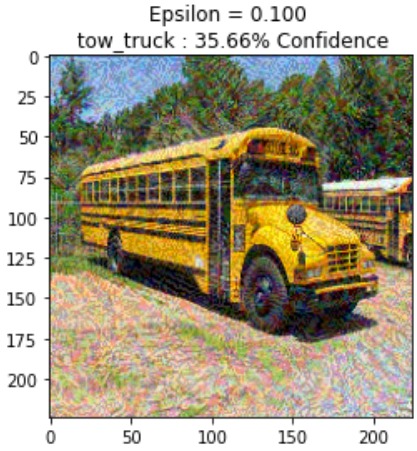
Predictive Score:

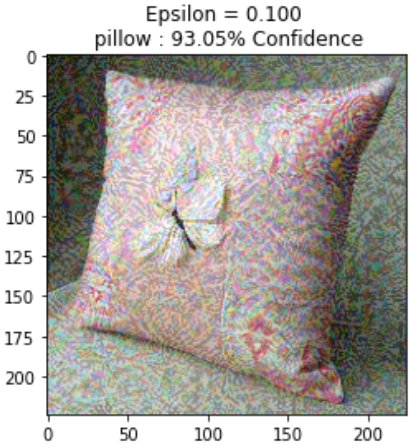


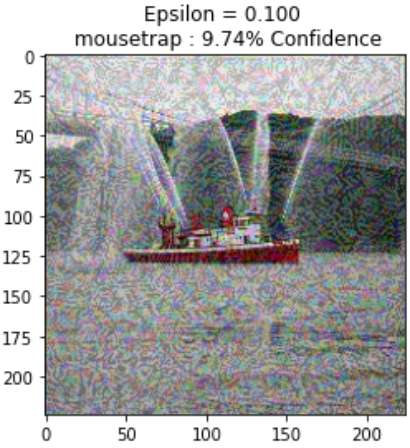


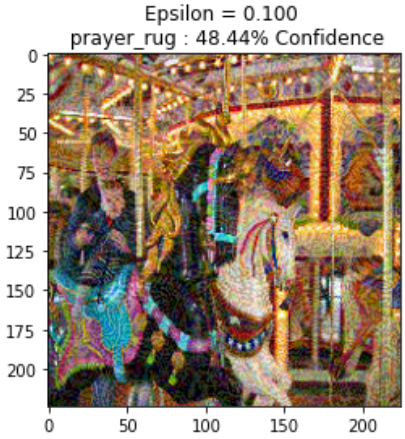












Analysis

The model does give another wrong label predictions on some images such as rapeseed(with even higher probability) and school bus. However, there are still images are predicted correctly with lower confidence, such as pillow and hourglass. Consider that the original images have extremely high confidence on classifying, it is hard to create adversarial examples.