图片：ImageNet中不同标签的1000张图片

模型：mobilenet\_v2，lama

1. 生成advImage，并分类

Epsilon: 0 Test Accuracy = 1.0

Epsilon: 0.01 Test Accuracy = 0.04

Epsilon: 0.02 Test Accuracy = 0.016

Epsilon: 0.03 Test Accuracy = 0.006

Epsilon: 0.04 Test Accuracy = 0.008

Epsilon: 0.05 Test Accuracy = 0.01

Epsilon: 0.06 Test Accuracy = 0.008

1. 选取e=0.01的生成的500个对抗样本

acc, ex = test(model, test\_loader, device, epsilon, test\_count=500)

1. 用Random\_medium\_256生成随机mask(detectron2)

python bin/gen\_mask\_dataset.py configs/data\_gen/random\_medium\_256.yaml $(pwd)/advImage $(pwd)/advOutput --ext JPEG

1. 然后用big-lama模型，修复mask后的advImage，得到

python3 bin/predict.py model.path=$(pwd)/big-lama indir=$(pwd)/advOutput outdir=$(pwd)/advInpainting

1. 用mobilenet\_v2分类修复后的图片

python class\_predict\_adv.py

Random\_medium\_256：Test Accuracy = 180 / 500 = 0.36

random\_thin\_256：Test Accuracy = 190 / 500 = 0.38

Random\_thick\_256：Test Accuracy = 166 / 500 = 0.332

random\_medium\_512：Test Accuracy = 161 / 500 = 0.322

random\_thick\_512：Test Accuracy = 140 / 500 = 0.28

random\_thin\_512：Test Accuracy = 178 / 500 = 0.356

Ps:由于lama中的文件匹配机制，图片名称中不能包含mask，否则将造成混乱

Epsilon: 0.01 Test Accuracy = 20 / 10000 = 0.002

random\_medium\_256：

random\_thin\_256：

random\_thick\_256：

random\_medium\_512：

random\_thick\_512：

Epsilon: 0.01 Test Accuracy = 409 / 977 = 0.4186284544524053

Test Accuracy = 34 / 971 = 0.035015447991761074

Epsilon: 0.02 Test Accuracy = 387 / 977 = 0.3961105424769703