

Method	White Box Top 1	White Box Top 5
Regular training	0.7%	4.4 %
Tramèr et al. (2018)	1.3%	6.5 %
Kurakin et al. (2017a)	1.5%	5.5 %
M-PGD	3.9%	10.3%
ALP	<b>27.9%</b>	<b>55.4%</b>