

Classifier/ Substitute	No Attack	No Defense	Defense- GAN-Rec	Defense- GAN-Orig	MagNet	Adv. Tr. $\epsilon = 0.3$	Adv. Tr. $\epsilon = 0.15$
A/B	0.9346	0.5131	0.586	0.5803	0.5404	0.7393	<u>0.6600</u>
A/E	0.9346	0.3653	0.4790	0.4616	0.3311	0.6945	<u>0.5638</u>
B/B	0.7470	0.4017	<u>0.4940</u>	0.5530	0.3812	0.3177	0.3560
B/E	0.7470	0.3123	<u>0.3720</u>	0.4187	0.3119	0.2617	0.2453
C/B	0.9334	0.2635	0.5289	0.6079	0.4664	0.7791	<u>0.6838</u>
C/E	0.9334	0.2066	0.4871	0.4625	0.3016	0.7504	<u>0.6655</u>
D/B	0.8923	0.4541	0.5779	0.5853	0.5478	<u>0.6172</u>	0.6395
D/E	0.8923	0.2543	0.4007	0.4730	0.3396	0.5093	<u>0.4962</u>