NEW EAD Attack against BSD at 2.0% Detection Threshold CIFAR10 100% 80% Classification accuracy 60% 40% MagNet 20% BSD with JSD BSD w/o JSD 0% 20 60 40 80

Confidence in EAD attack