

1. Original image

Logo



Blank Sign

Traffic Sign

**Physically robust
adversarial example**



Logo
attack



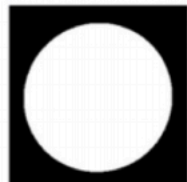
Custom
Sign attack



Adversarial
Traffic Sign

Attack Pipeline

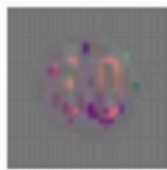
2.a Find mask to limit
adversarial perturbation
to sign areas (Canny edge
detection + Fill holes)



**2.b Resize original
image and mask** to
match the neural
network's input size



**2.d Resize perturbation
and add to original image**
to create a high-res
printable fake sign



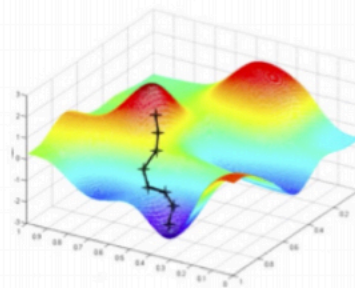
Original image

2.c Optimization with random transformation

Optimization Output
(low-res adv.
examples)



Find optimal
perturbation with
Adam optimizer



Batch of randomly
transformed samples



...