

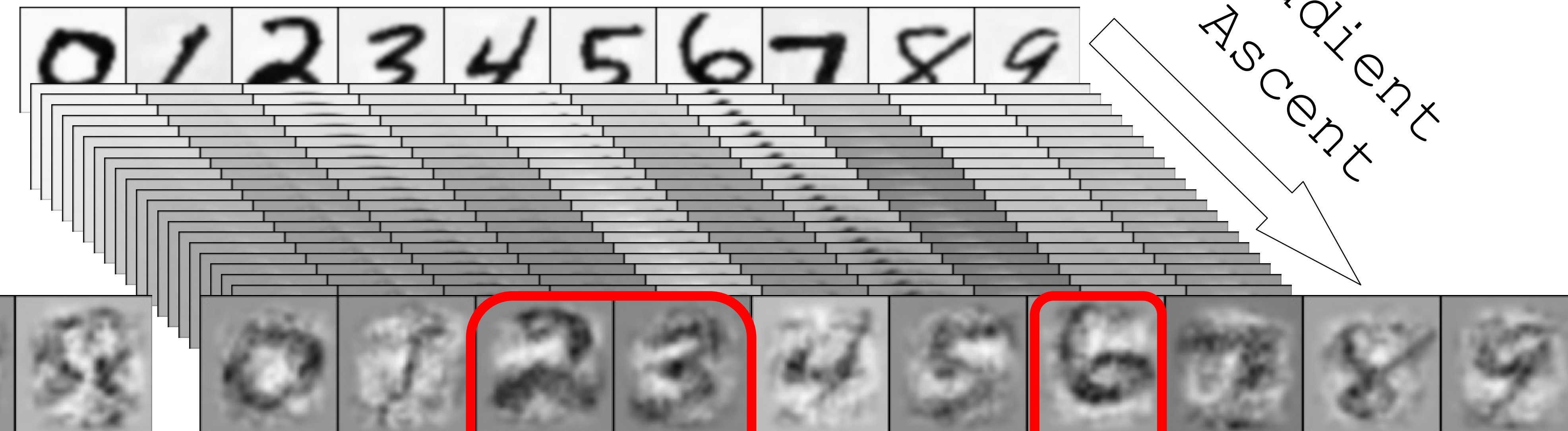
Adversarial samples

Visualization candidates

Gradient  
Ascent

Biased

$[-1, -1 \dots 1 \dots -1, -1]$



Unbiased

$[0, 0 \dots 1 \dots 0, 0]$

