



$+ .007 \times$



$=$



$\mathbf{x}$

$\text{sign}(\nabla_{\mathbf{x}} J(\boldsymbol{\theta}, \mathbf{x}, y))$

$\mathbf{x} + \epsilon \text{sign}(\nabla_{\mathbf{x}} J(\boldsymbol{\theta}, \mathbf{x}, y))$

$y = \text{"panda"}$

$\text{"nematode"}$

$\text{"gibbon"}$

w/ 57.7%

w/ 8.2%

w/ 99.3 %

confidence

confidence

confidence