Normalization

Normalization

- How to prevent vanishing (or exploding) gradients?

x

Conv	Conv
ReLU	ReLU
Conv	Conv
ReLU	ReLU
Conv	Conv
⋮	⋮
Conv	Conv

$$\frac{\partial \ell(\mathbf{o})}{\partial \mathbf{o}}$$

o