

Failed defenses

Generative
pretraining

Removing perturbation
with an autoencoder

Adding noise
at test time

Ensembles

Confidence-reducing
perturbation at test time

Error correcting
codes

Multiple glimpses

Weight decay

Double backprop

Adding noise
at train time

Various
non-linear units

Dropout