

ImageNet-trained CNNs are biased towards texture:

Increasing shape bias improves accuracy and robustness

shape hypothesis

texture hypothesis

A cat with elephant feature → CNNs: elephant

Humans: cat



Bias can be overcome w/ proper dataset

Stylized-ImageNet (sIN)

: strip every single image of original texture and replace with the style of a randomly selected painting through style transfer.

ImageNet-trained CNNs tend to classify objects according to local textures instead of global object shapes.

work poorly on image distortions

generalize better to natural images.

shape: relatively stable

texture: easily distorted by all sorts of noise