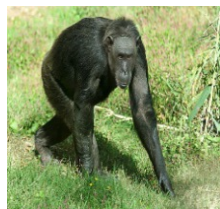


original sample  
label: dog



original sample  
label: cat



original sample  
label: primate



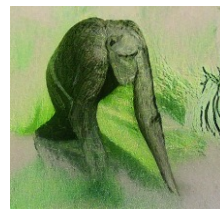
unbounded  
perturbation



valid perturbation  
predicted: dog



valid perturbation  
predicted: bird



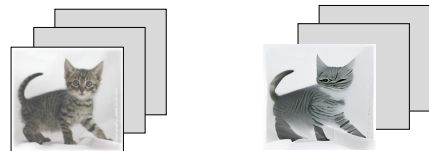
invalid perturbation

finite dataset for specific application  
and trained models to evaluate

massive generic data and pretrained models

real-world data space

a. main structure of proposed evaluation method



c. sparsify latent representations



encoder



decoder



b. unbounded perturbation with sparse VQGAN for style-only perturbation