Deep Generative Models for Anomaly Detection

Lawrence Carin
Duke University

Motivation

- In many DOE sensing scenarios, we may have knowledge of what “typical” looks like, but atypical/anomalous may be manifested in unanticipated ways

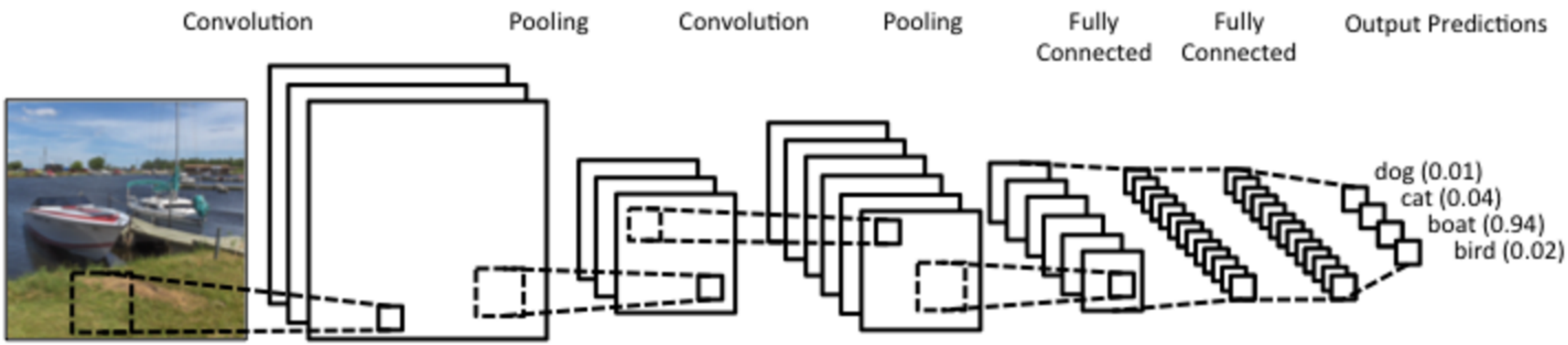
- Desire statistical model of typical data, with ability to compute the likelihood that new data under test matches the statistical model

- A classic problem, but traditional methods (maximum likelihood) fail

- There has recently been a “revolution” in the ability to learn generative statistical models from which highly realistic (typical) data may be drawn

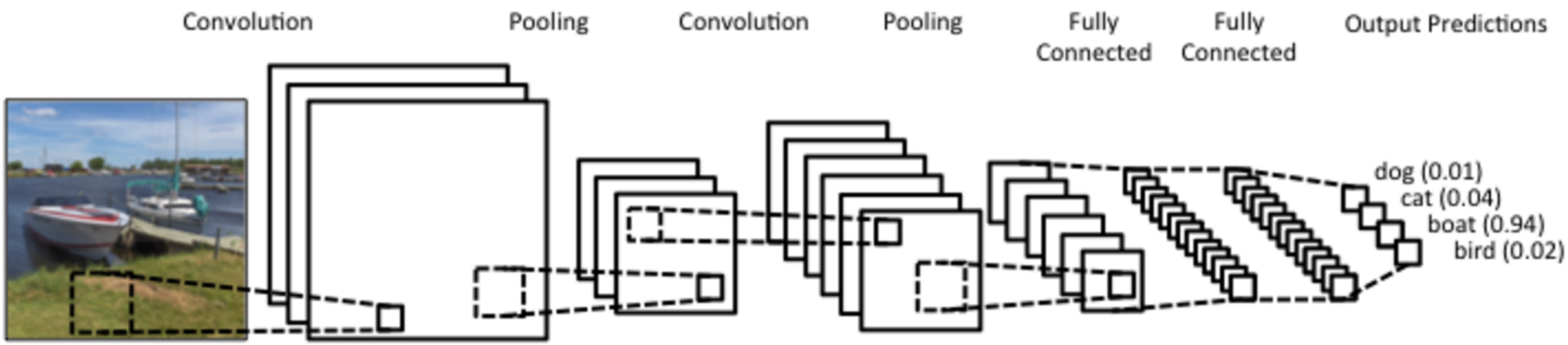
- Extend those such that we may use such models to compute the likelihood that new test data match the learned generative statistical model

Traditional Deep Learning: Feed-Forward



- Excellent model for performing classification with *known* labels
- Requires large quantity of labeled data for all targets of interest
- All training images must be labeled
- Less appropriate for many DOE monitoring applications

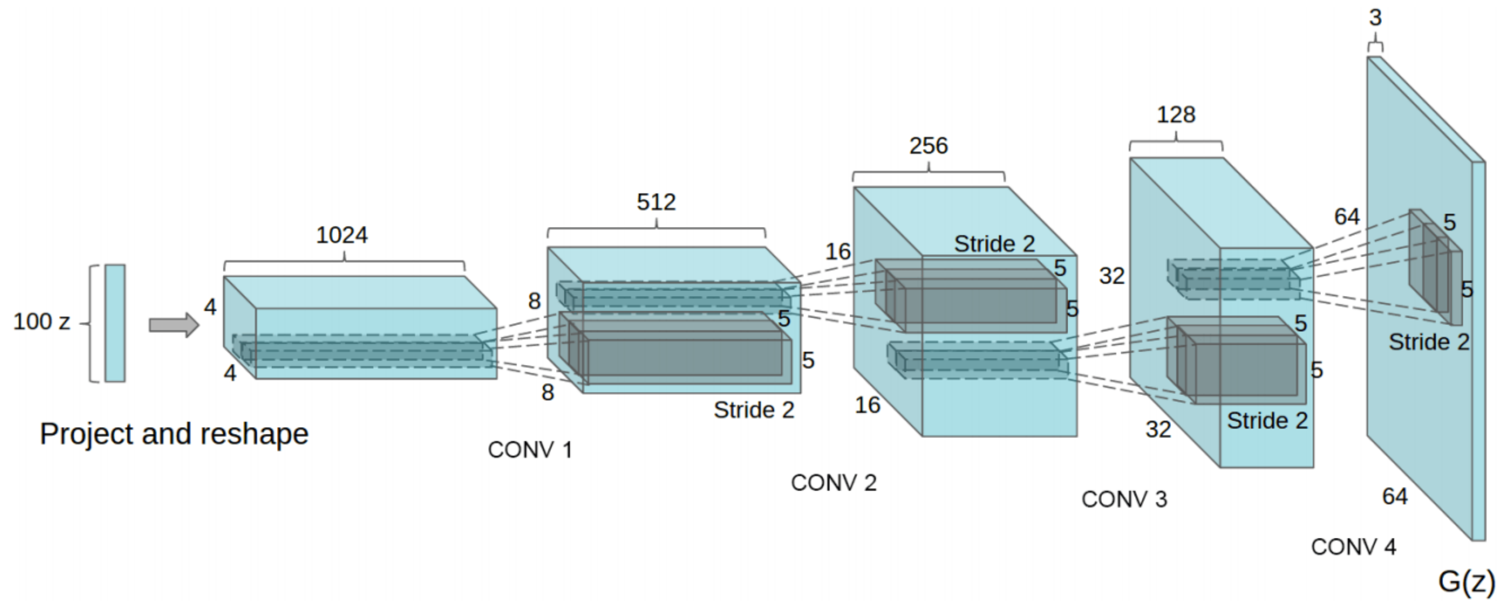
Traditional Deep Learning: Feed-Forward



- Excellent model for performing classification with *known* labels
- Requires large quantity of labeled data for all targets of interest
- All training images must be labeled
- Less appropriate for many DOE monitoring applications

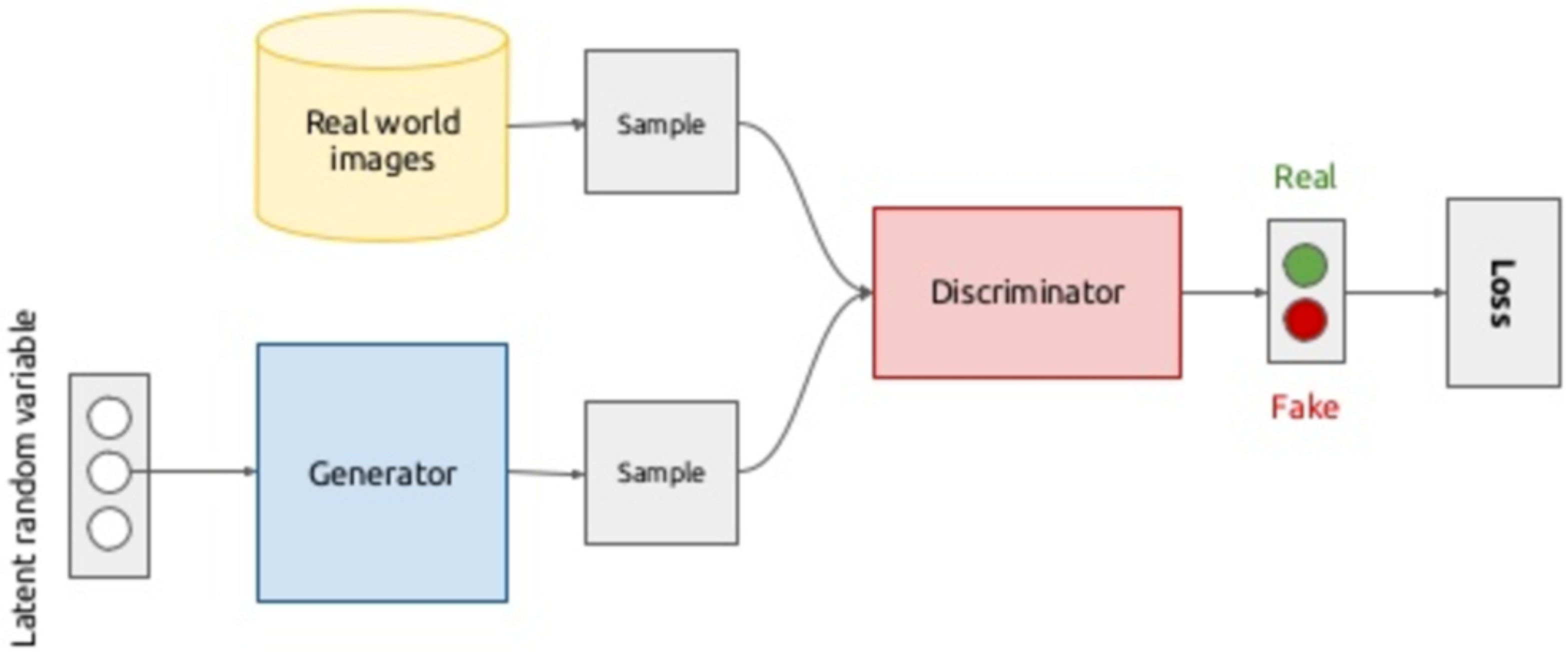
  - ❖ Doesn't leverage vast quantities of unlabeled data
  - ❖ Not appropriate for previously unseen targets ("black swan")

Generative Model: From Noise to Image



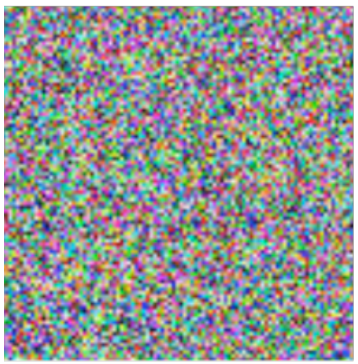
- Vector **z** drawn from simple distribution (e.g., isotropic Gaussian)
- Via sequence of deconvolutions and nonlinearities, generate an image
- Transform RVs **z** drawn from a simple distribution to RVs drawn from distribution for the data of interest, with the nonlinear functional transformation learned

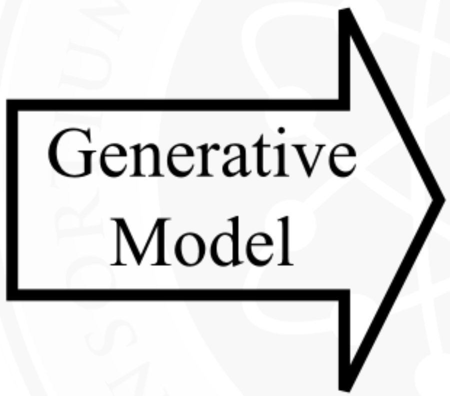
Training with Unlabeled Images

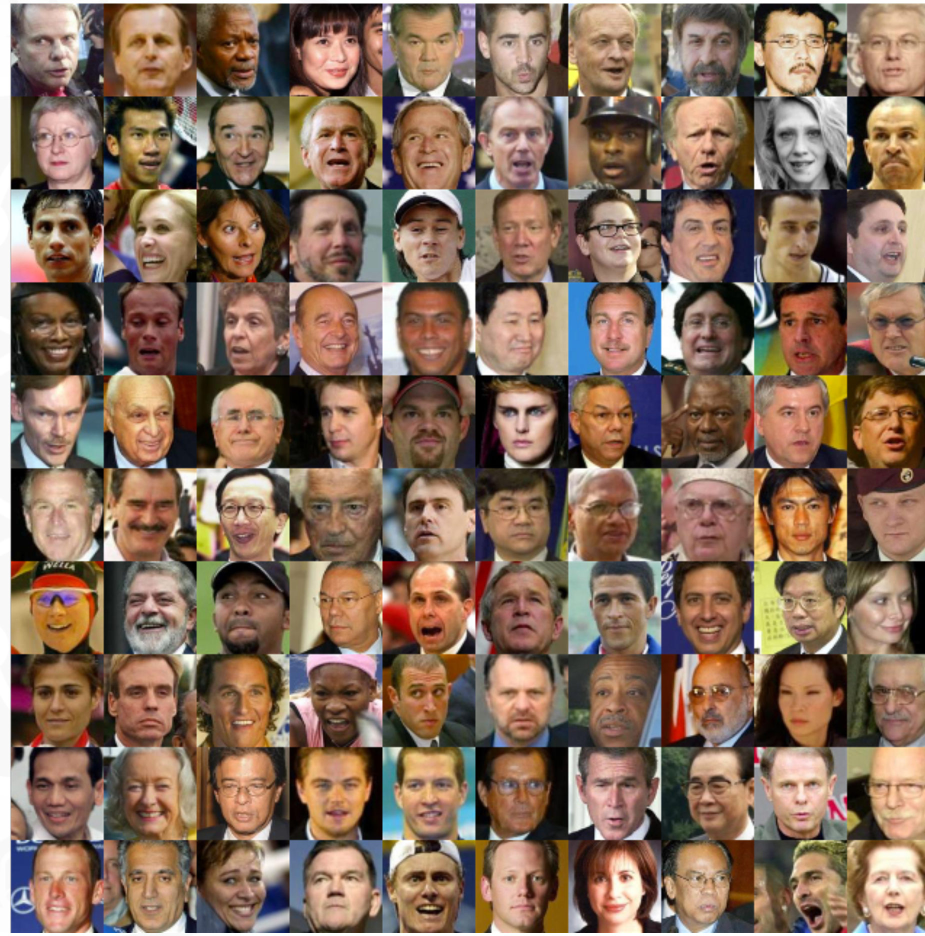


Example Images

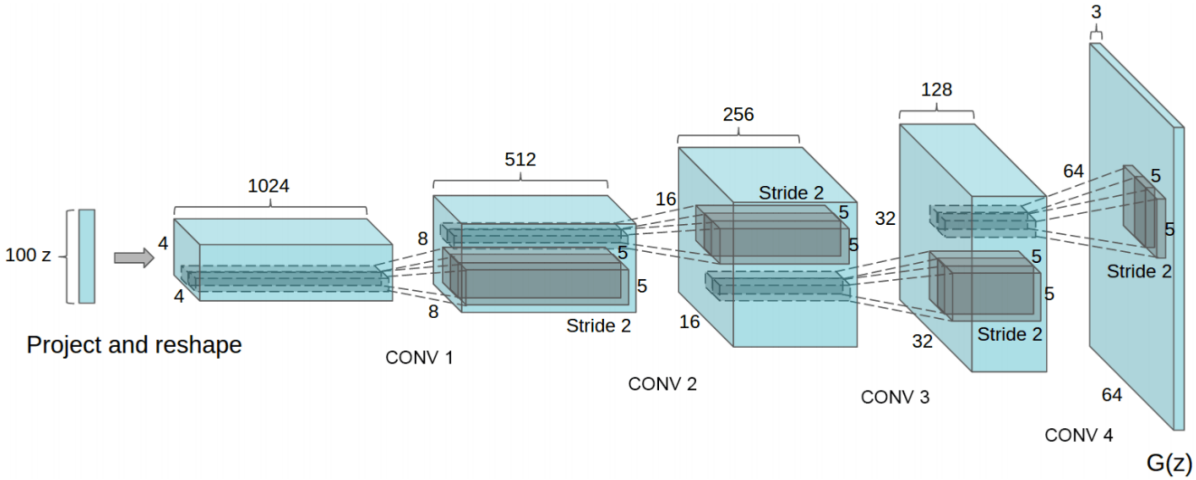
Noise ~ N(0,1)





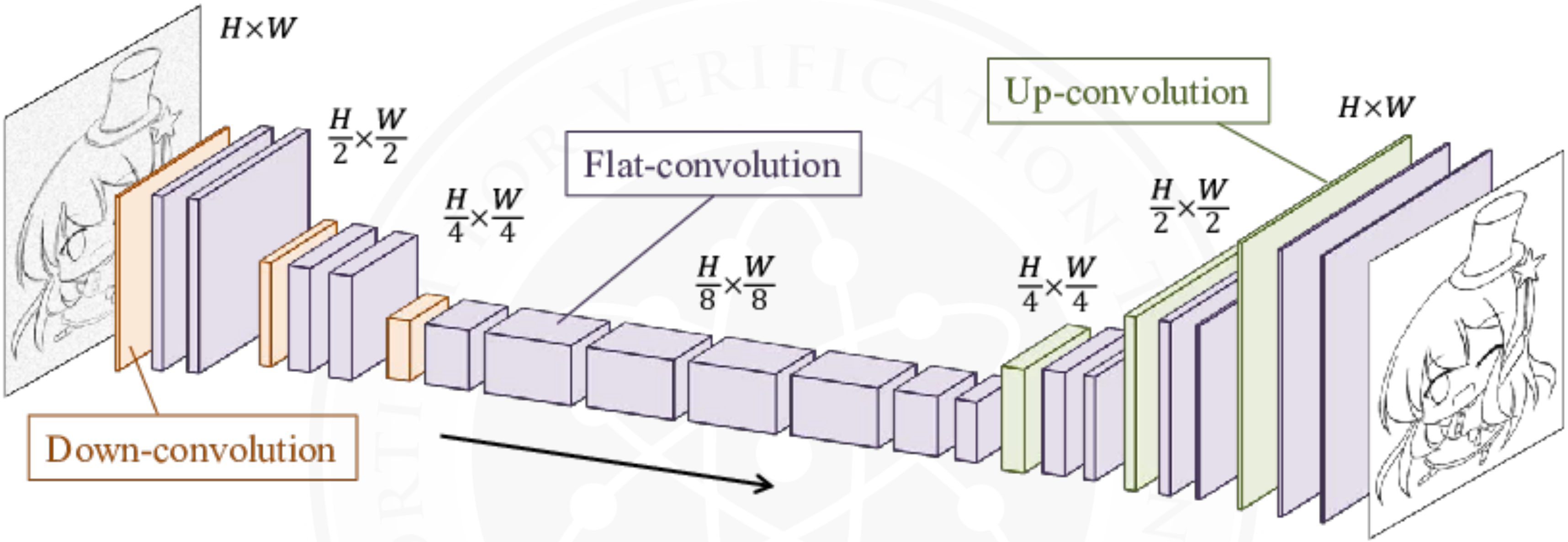


Problem



- We have a model capable of synthesizing realistic images
- Implemented by drawing from simple distribution, and then transforming drawn RVs, via deep nonlinear functional operation
- We do not have an explicit functional relationship for statistical distribution of data
- Cannot assess whether new data are consistent with the distribution, and this is needed for anomaly detection

Convolutional Autoencoder



- The encoder "arm" of this model is like the supervised classifier
- The decoder arm of this model is like the generative model
- Put these two together: can train on labeled and unlabeled data

Symmetric Variational Autoencoder

- A direct implementation of the convolutional autoencoder does *not* yield a model that matches the statistics of the data of interest

- This is connected to fundamental limitations in learning a generative statistical model via maximum-likelihood learning

- Have developed a new symmetric convolutional variational autoencoder

- Model synthesizes highly realistic data

- Also allows inference on test data, quantififying fit of test images to learned statistical model

Synthetic Images, ImageNet Training

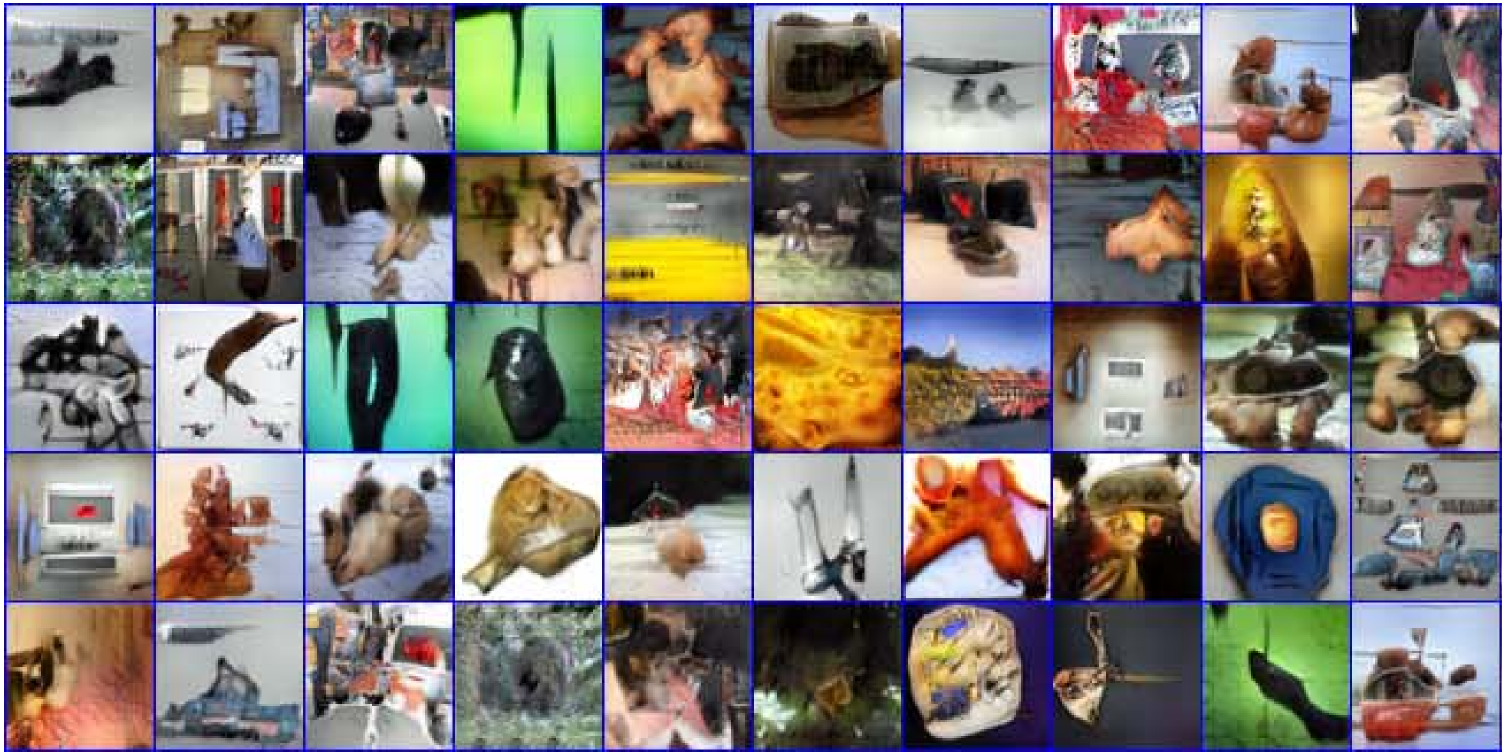


Image Transformation

Synthesis & Adjusted Attributes

Synthesis and Attributes

Summary

- In many DOE sensing scenarios, we may have knowledge of what “typical” looks like, but atypical/anomalous may be manifested in unanticipated ways

- Desire statistical model of typical data, with ability to compute the likelihood that new data under test matches the statistical model

- A classic problem, but traditional methods (maximum likelihood) fail

- There has recently been a “revolution” in the ability to learn generative statistical models from which highly realistic (typical) data may be drawn

- Extend those such that we may use such models to compute the likelihood that new test data match the learned generative statistical model