Generative Models

Image Generation

- Large outputs
- Small inputs
- Many possibilities

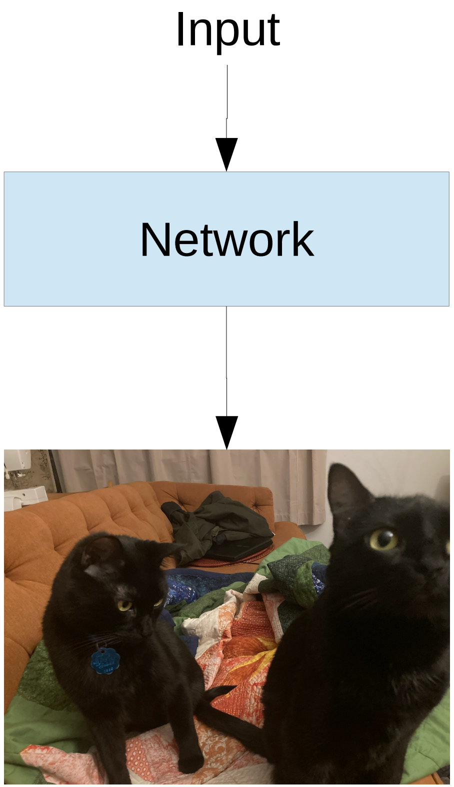
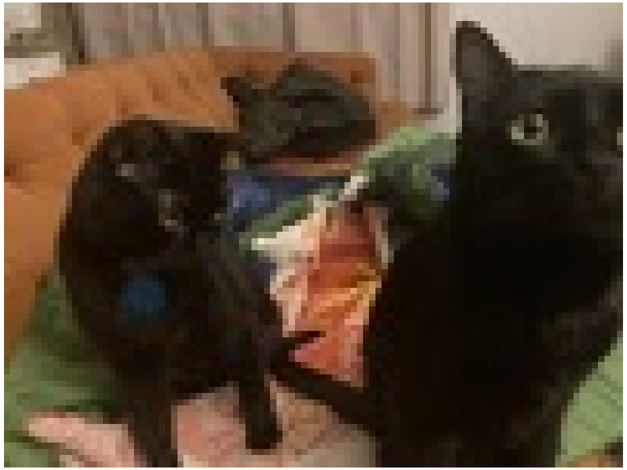


Image Editing

Denoising

Inpainting

Super-resolution

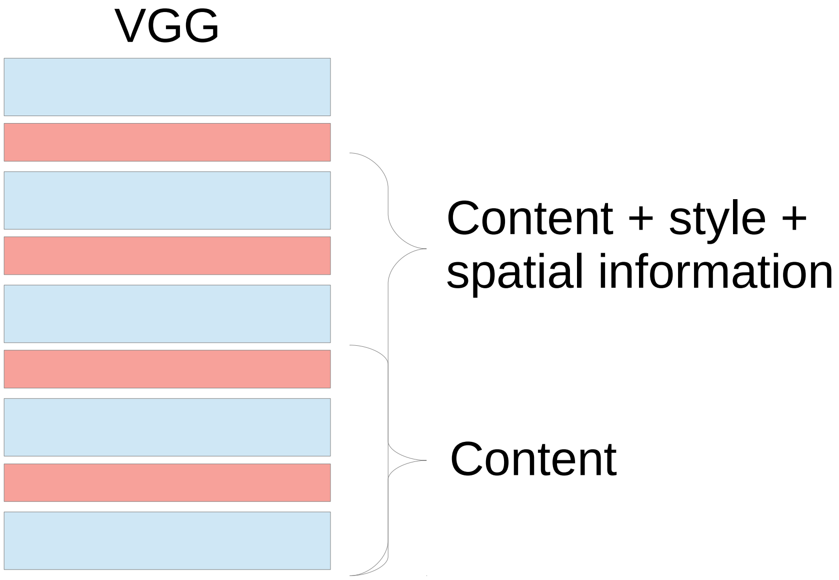


Qifeng Chen, Jia Xu, Vladlen Koltun. "Fast Image Processing with Fully-Convolutional Networks." ICCV 2017.

Style Transfer

Combine “content” from one image with “style” from another

Image from [1]

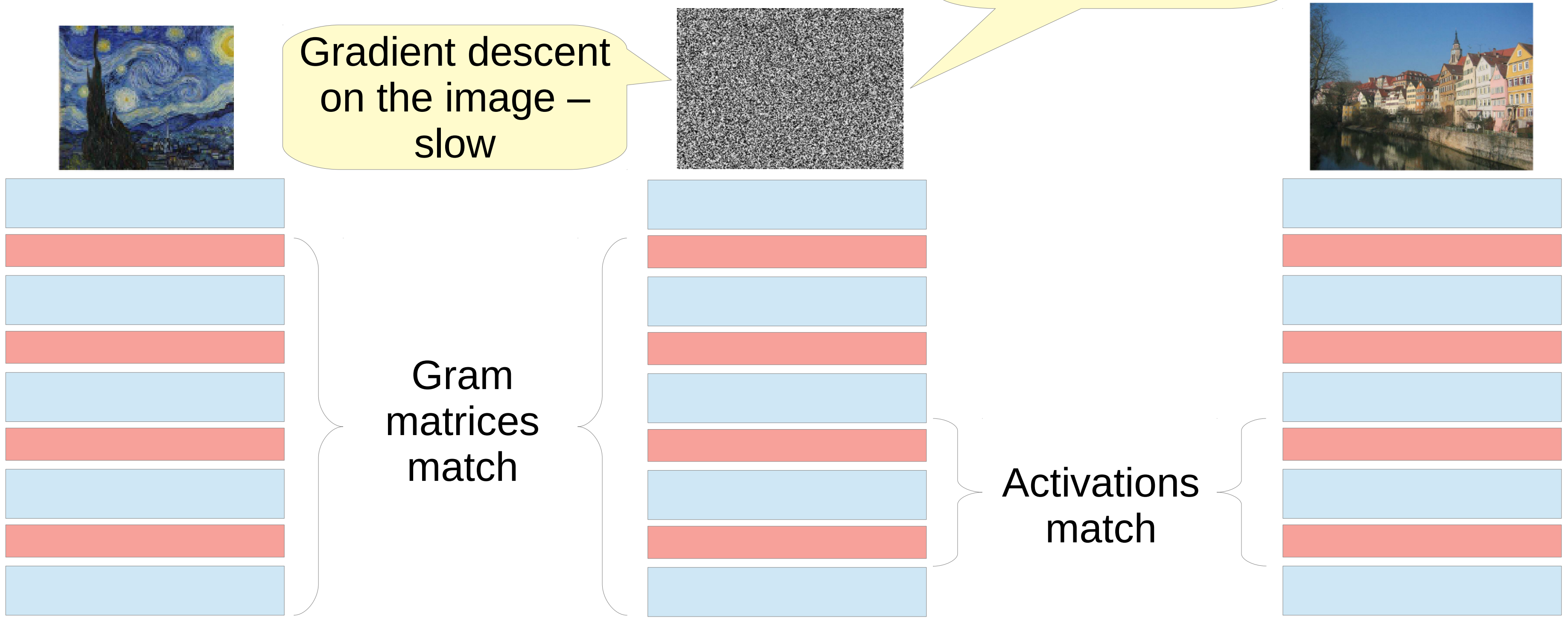


[1] Leon A. Gatys, Alexandre S. Ecker, Matthias Bethge. “Image Style Transfer Using Convolutional Neural Networks.” CVPR 2016.

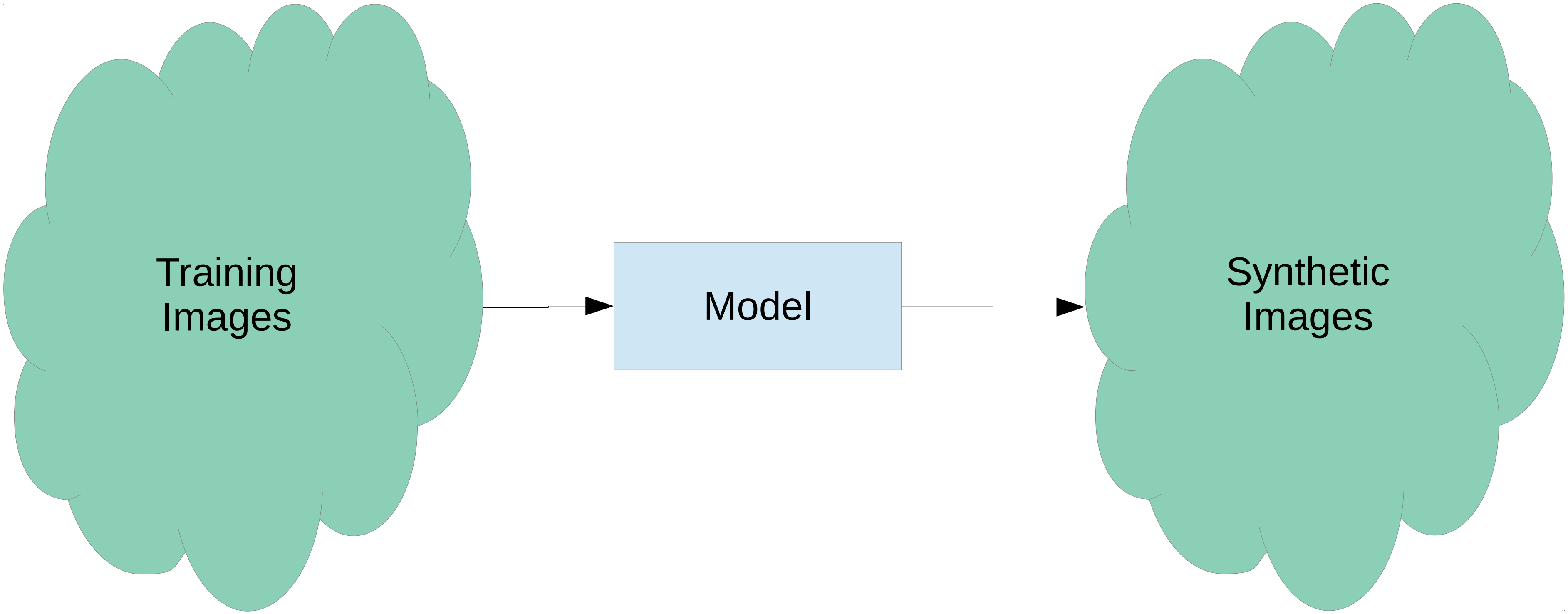
[2] Justin Johnson, Alexandre Alahi, Li Fei-Fei. “Perceptual Losses for Real-Time Style Transfer and Super-Resolution.” ECCV 2016.

Style Transfer

Gram Matrix: $G_{ij} = \sum_{w} \sum_{h} F_{iwh} F_{jwh}$



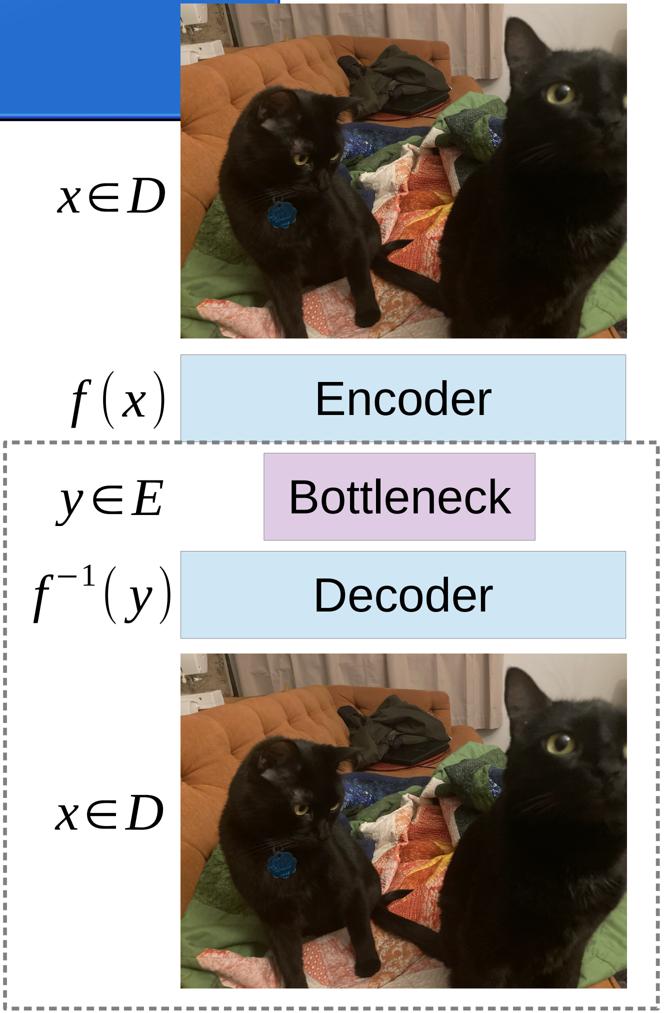
Sampling Images



Autoencoders

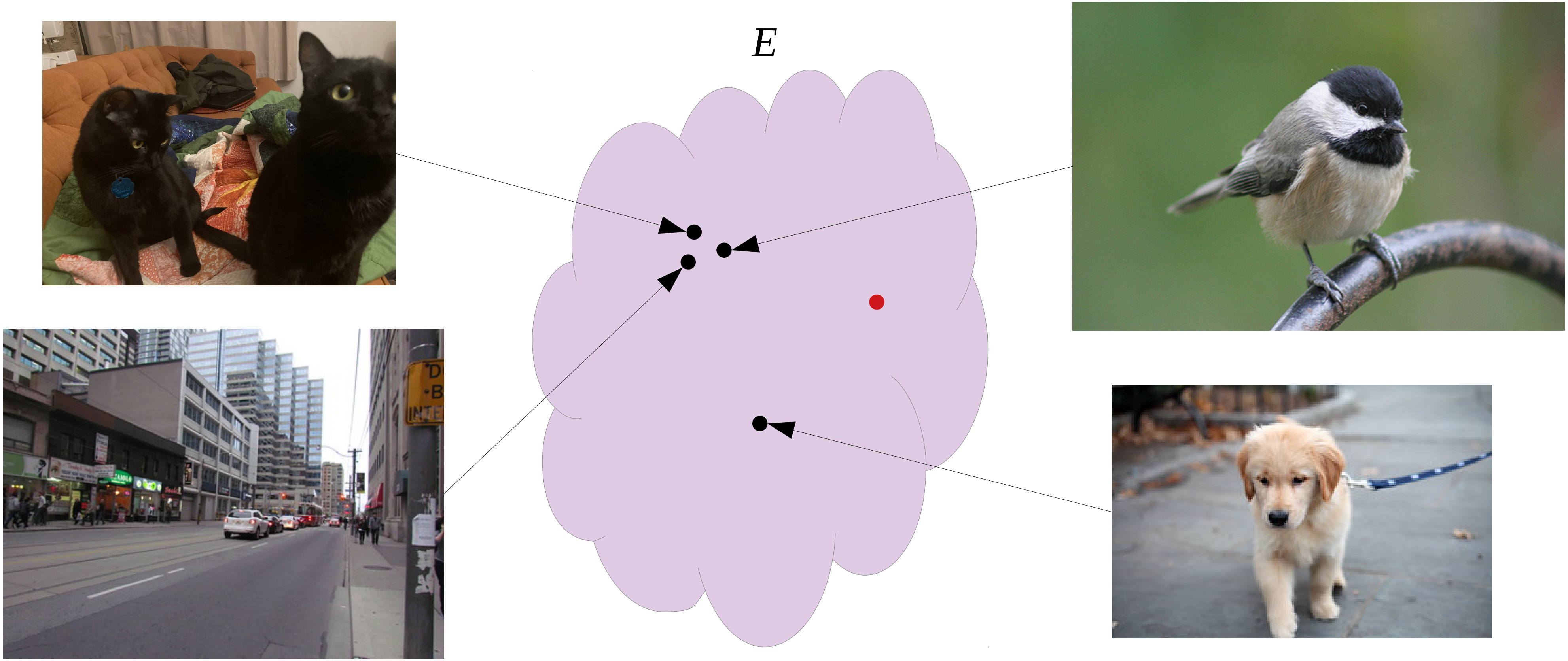
- Learn a low-dimensional representation of inputs

- Decoder as a generator

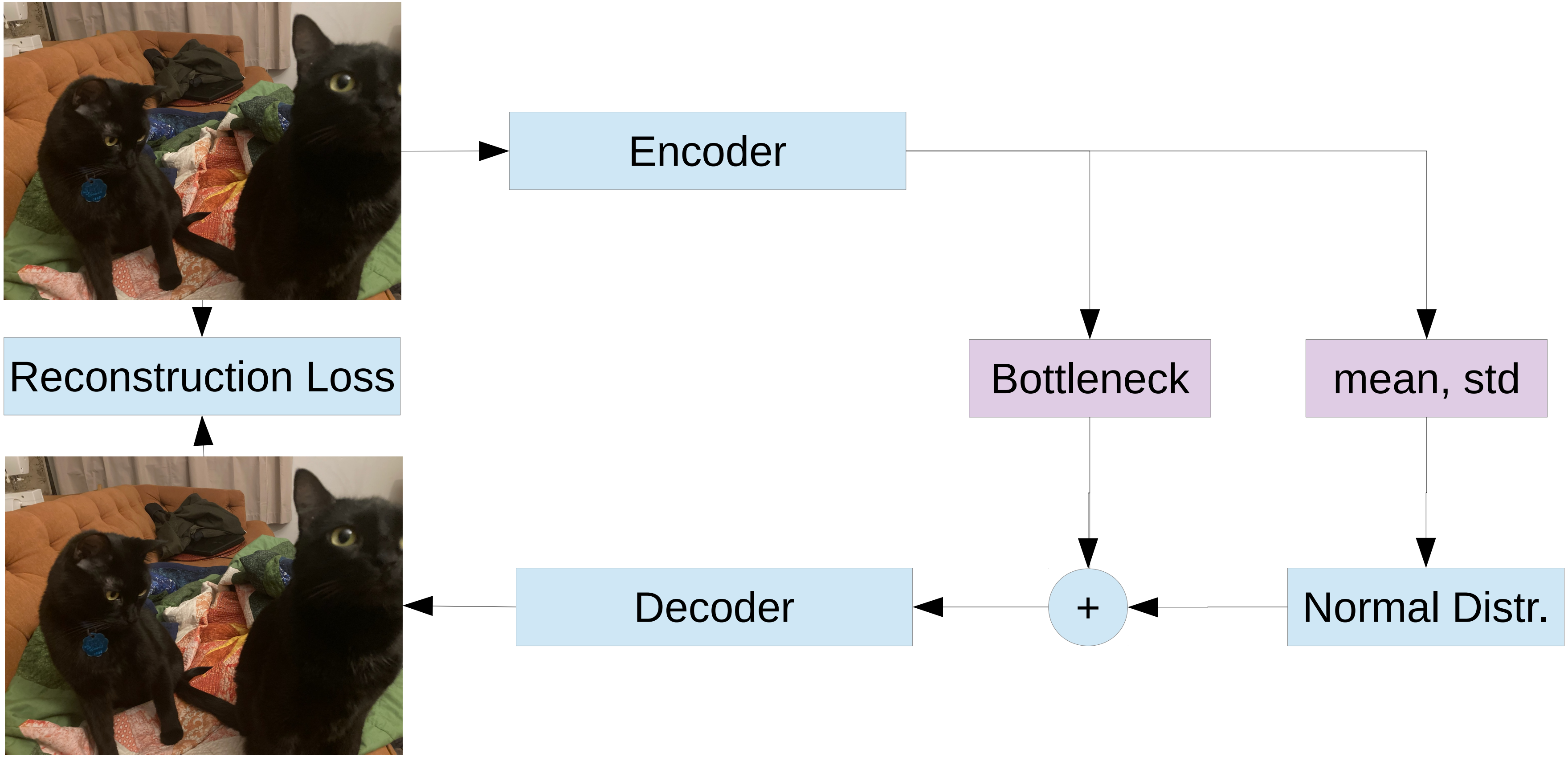


Mark A. Kramer. “Nonlinear Principle Component Analysis Using Autoassociative Neural Networks.” AIChE Journal. Feb. 1991.

Challenges with Autoencoders

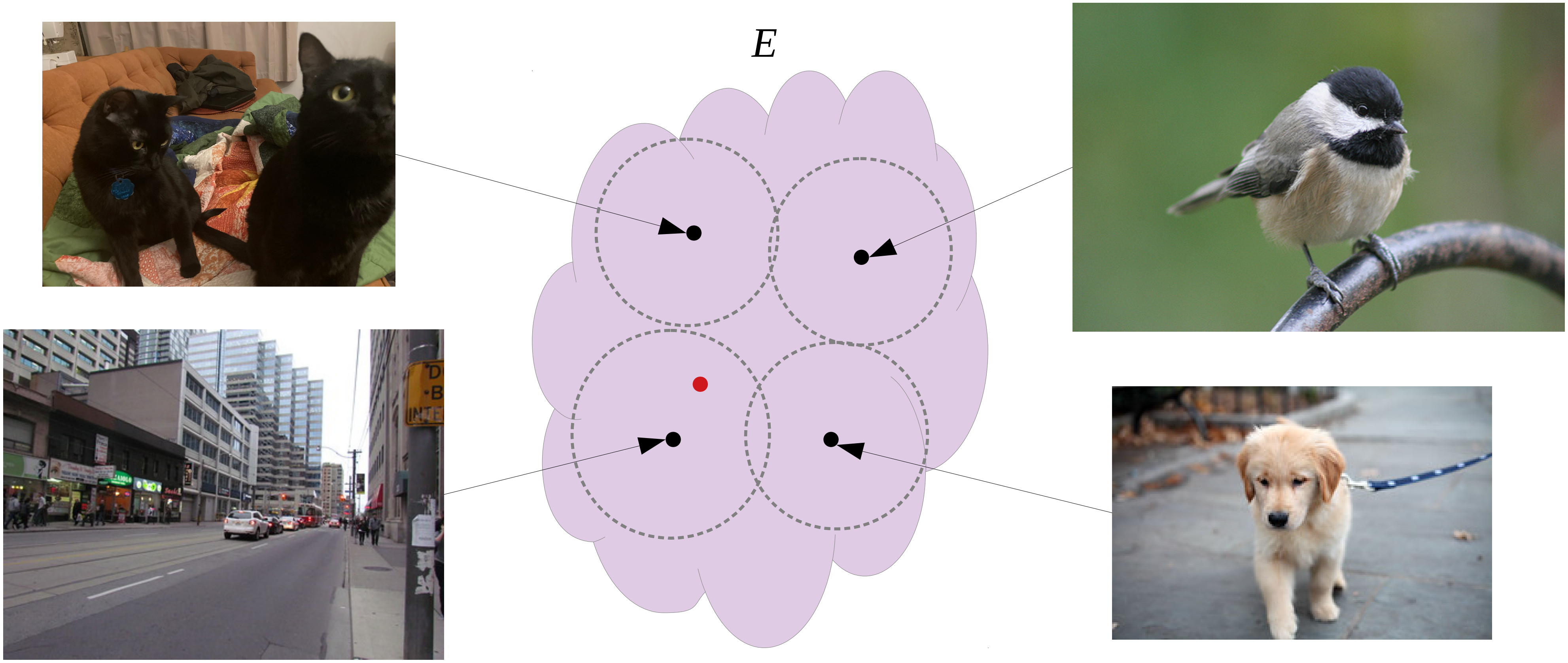


Variational Autoencoders (VAE's)



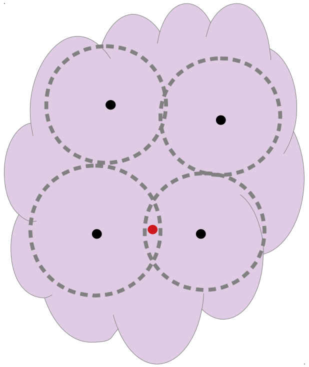
Diederik P Kingma, Max Welling. "Auto-Encoding Variational Bayes." ICLR 2014.

VAE Embeddings

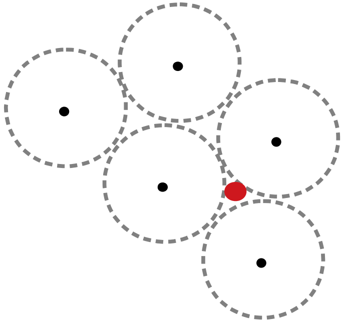


VAE Challenges

- Blurry Outputs



- High dimensions



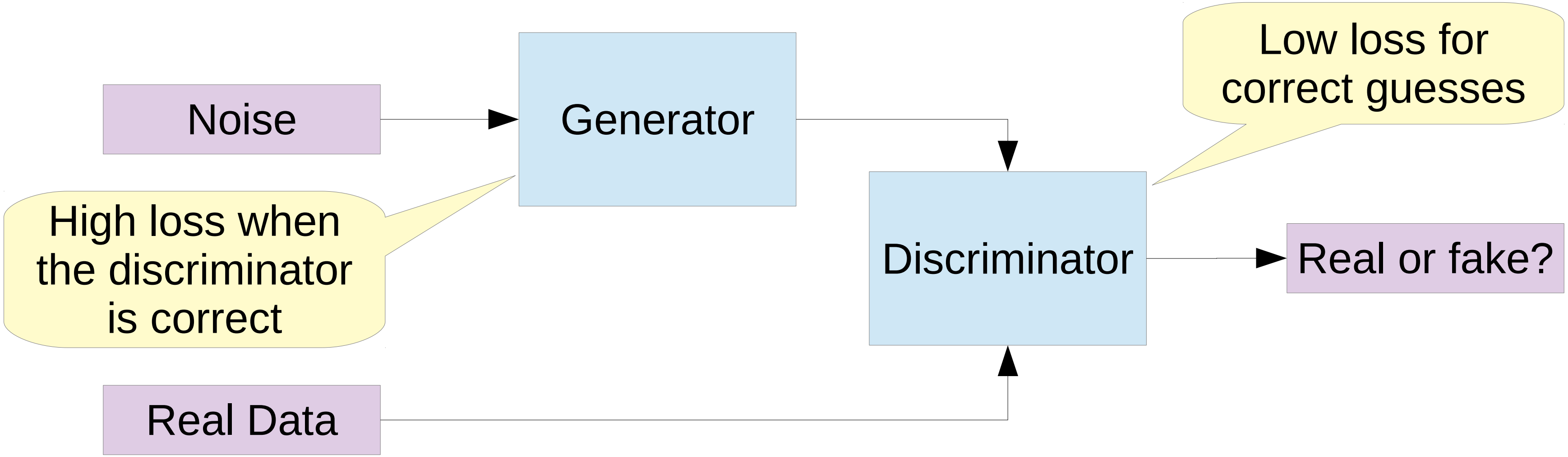
Transforming Noise

- To sample images, model the data distribution P_{Data}

Gaussian Noise → Neural Network → Data Distribution

- High-dimensional noise: no one-to-many issues
- Loss – how do we know if the network produces a good distribution?

Generative Adversarial Networks



$$\mathrm{argmin}_G \, \mathrm{max}_D \, D(\text{"fake"} \mid G(z))$$

$$P_{\text{Generator}} \approx P_{\text{Data}}$$

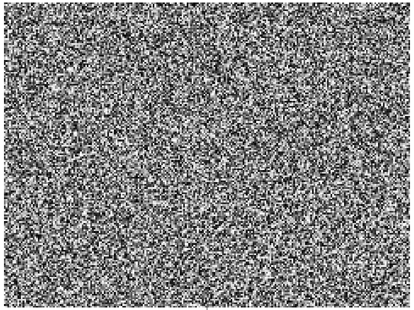
Jensen-Shannon Divergence, Wasserstein Metric

Ian Goodfellow, et al. Generative Adversarial Nets. NeurIPS 2014.

Applications

“Conditional GAN’s”

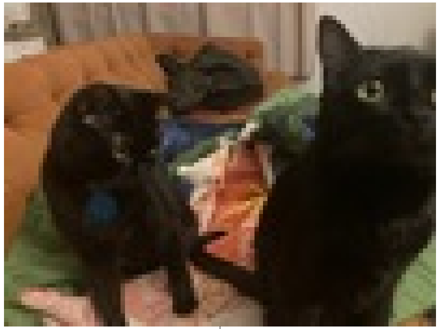
Sampling



Text-to-images

“Two black cats sitting on a quilt on an orange couch”

Super-resolution



Han Zhang, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, Dimitris Metaxas. “StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks.” ICCV 2017.

Christian Ledig, et al. “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network.” CVPR 2017.

pix2pix

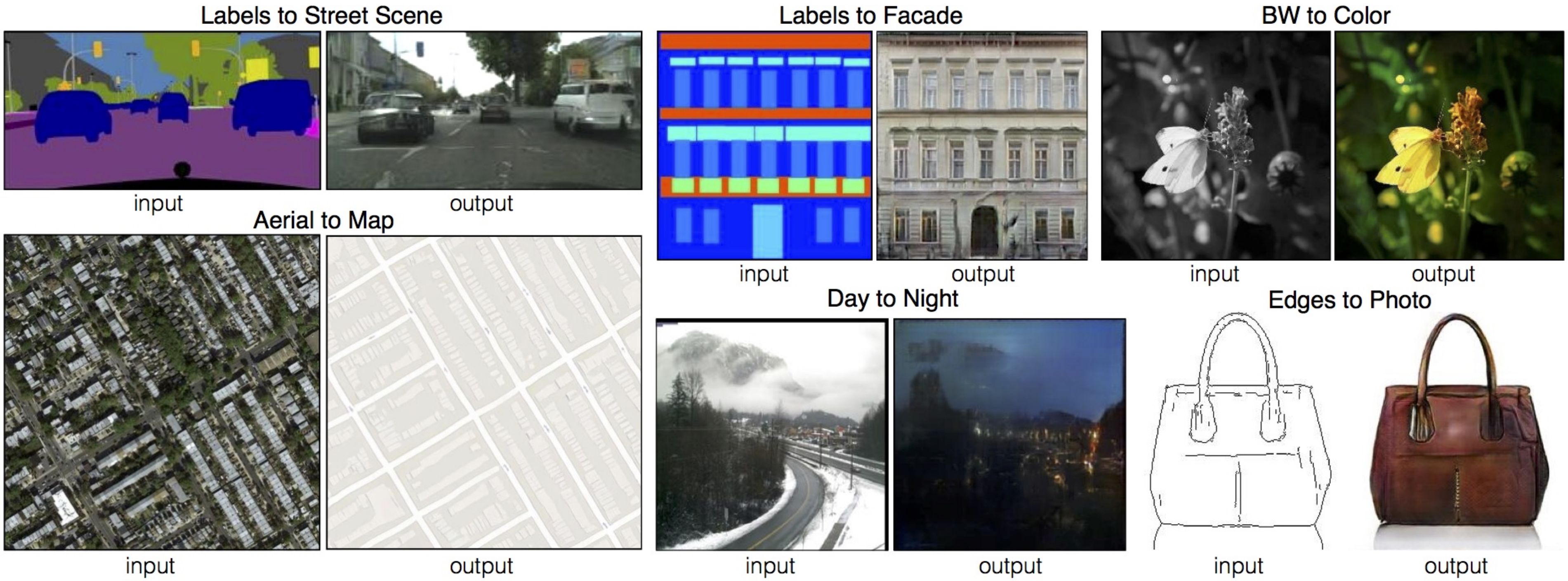
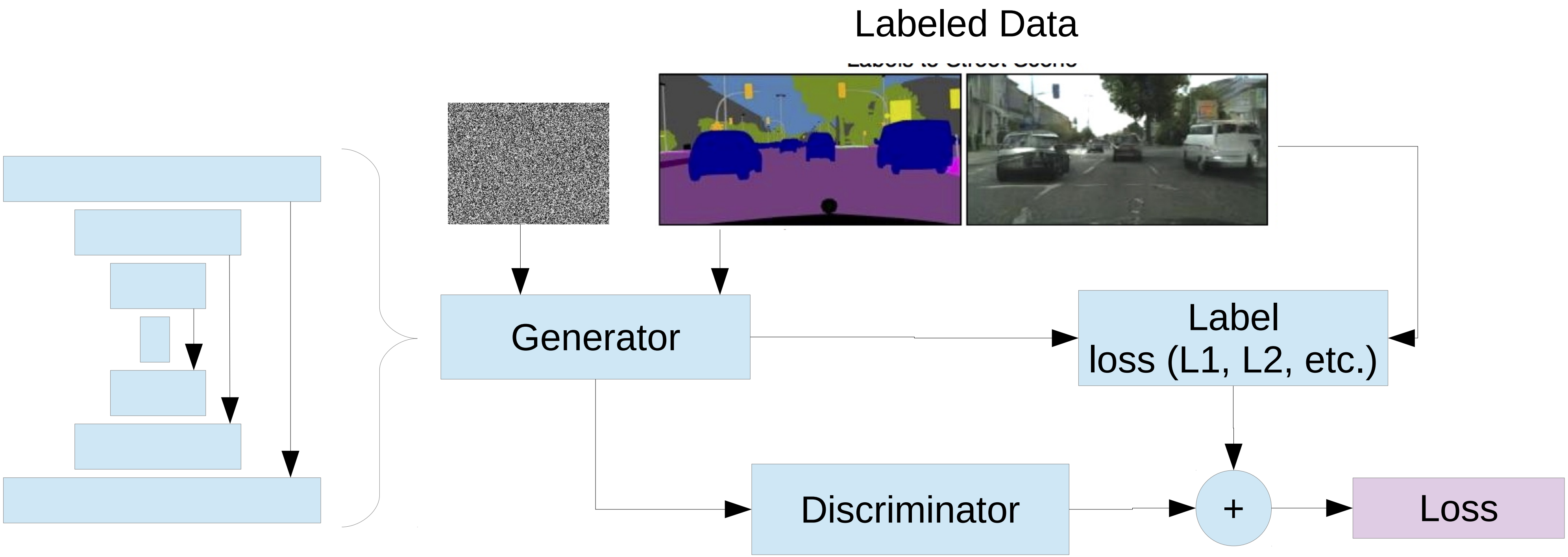


Image from: Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. “Image-to-Image Translation with Conditional Adversarial Networks.” CVPR 2017.

pix2pix



Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. “Image-to-Image Translation with Conditional Adversarial Networks.” CVPR 2017.

https://phillipi.github.io/pix2pix/

CycleGAN

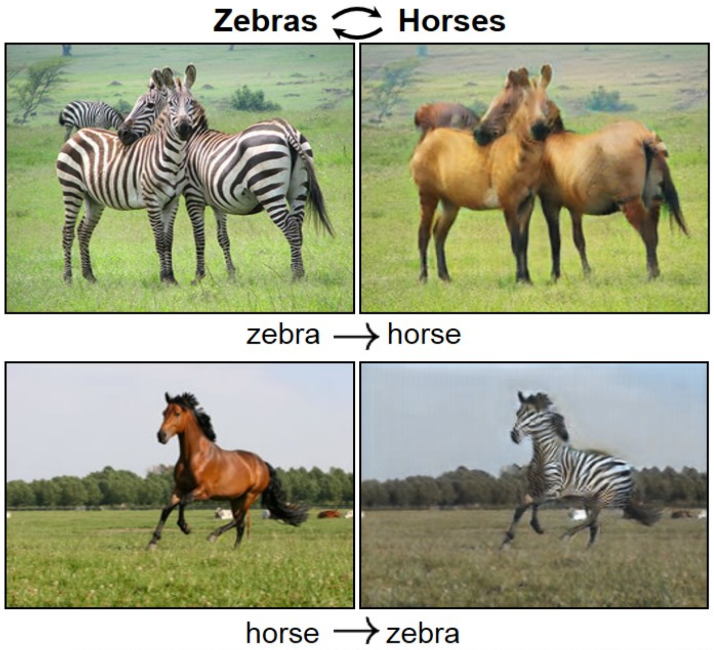
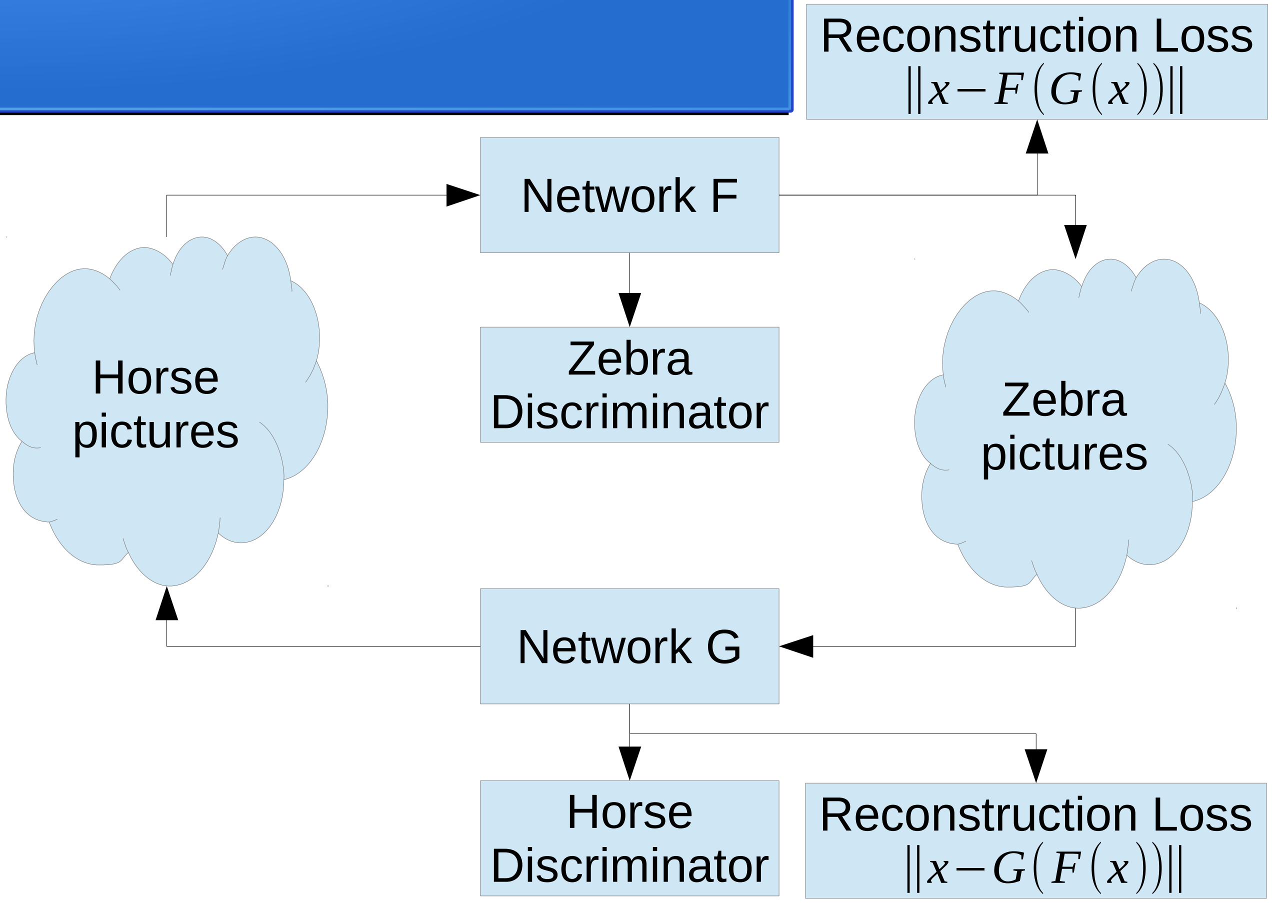


Image from [1]



[1] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros. “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.” CVPR 2017.

Data Augmentation

Train a GAN for each class to generate new images

- ✓ Provides more training data
- ✓ Free labels
- ✓ Sensitive/Unbalanced data

- ✗ Is the new data meaningful?
- ✗ In practice, other data augmentation seems to work better