Poor man’s survey on Generative Models

Sepehr Sameni
University of Tehran

What is generative modeling?[1]

Today I will use the following definition of a generative model:

A model is generative if it places a joint distribution over all observed dimensions of the data.

Generative versus discriminative supervised learning[1]

Consider a supervised learning task with features X and labels Y:

- Generative models want to learn P(X, Y).
- Discriminative models want to learn P(Y | X).

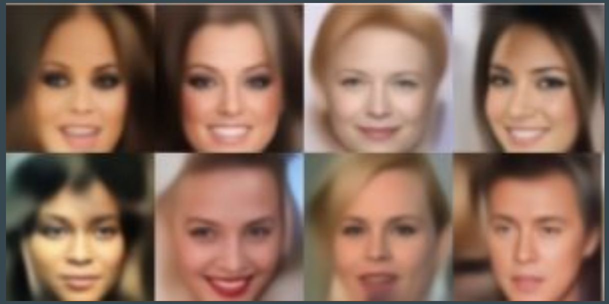
"… one should solve the [classification] problem directly and never solve a more general problem as an intermediate step … " Vapnik (1998)

- Sample efficiency[2]
- Unsupervised proxy
- Beyond P(Y | X)

Generative models: beyond P(Y | X)[1]

What can you do with a generative model?

- Compute arbitrary conditionals and marginals.
- Compare the probabilities of different examples.
- Reduce the dimensionality of the data.
- Identify interpretable latent structure.
- Fantasize completely new data.

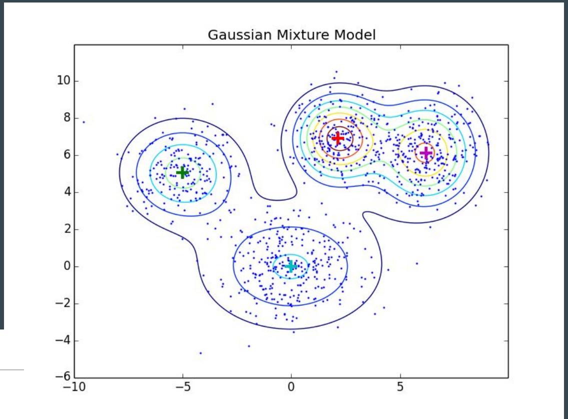


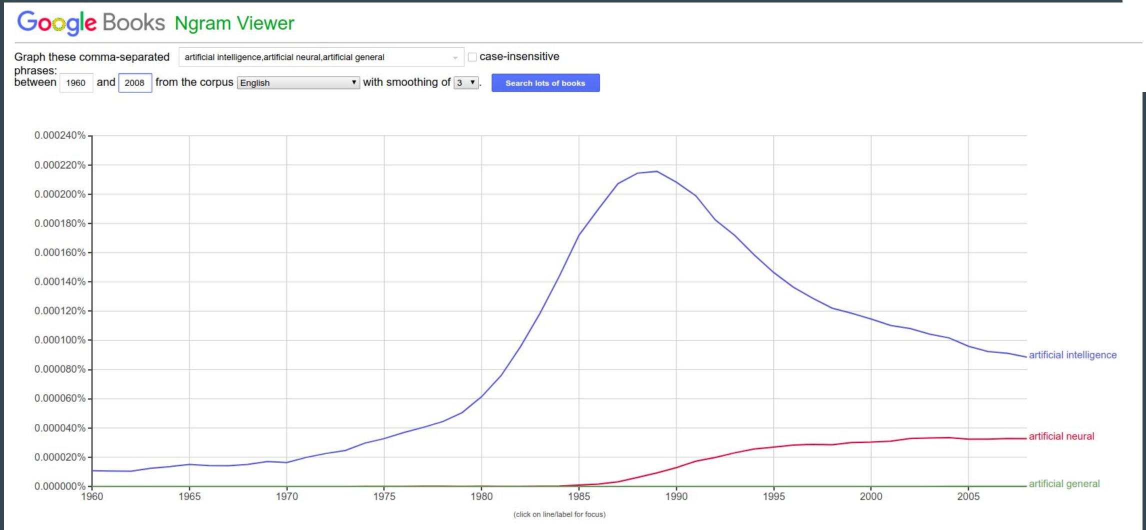
Generative Methods

- Old methods
- Autoregressive
- Variational methods
- Generative Adversarial Networks
- Flow based
- Energy based
- Mix & Match!

Old Methods

- GMM
- N-gram
- LDA





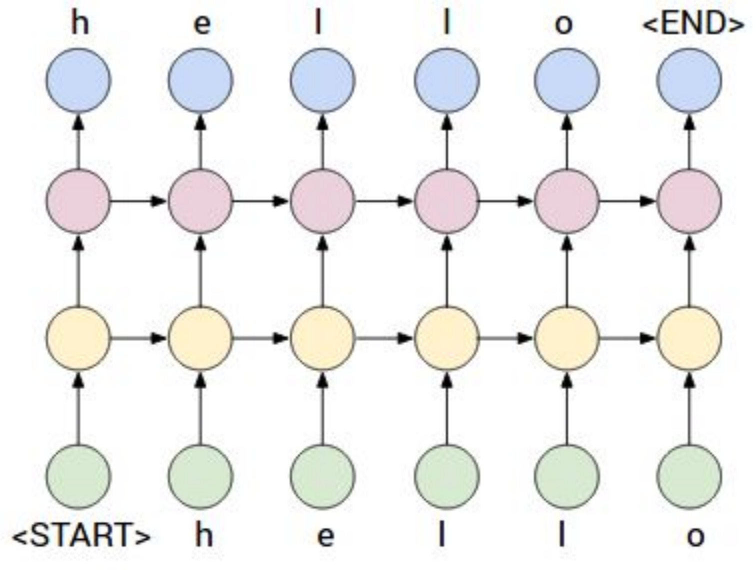
Autoregressive Models I

Autoregressive generative models over high-dimensional data $\mathbf{x} = (x_1, \ldots, x_n)$ factor the joint distribution as a product of conditionals:

$$p(\mathbf{x}) = p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i | x_1, \ldots, x_{i-1})$$

Autoregressive Models II (text)

Modern Language Modeling[3]



Autoregressive Models III (image)

- Pixel RNN[4]
- Pixel CNN[5]

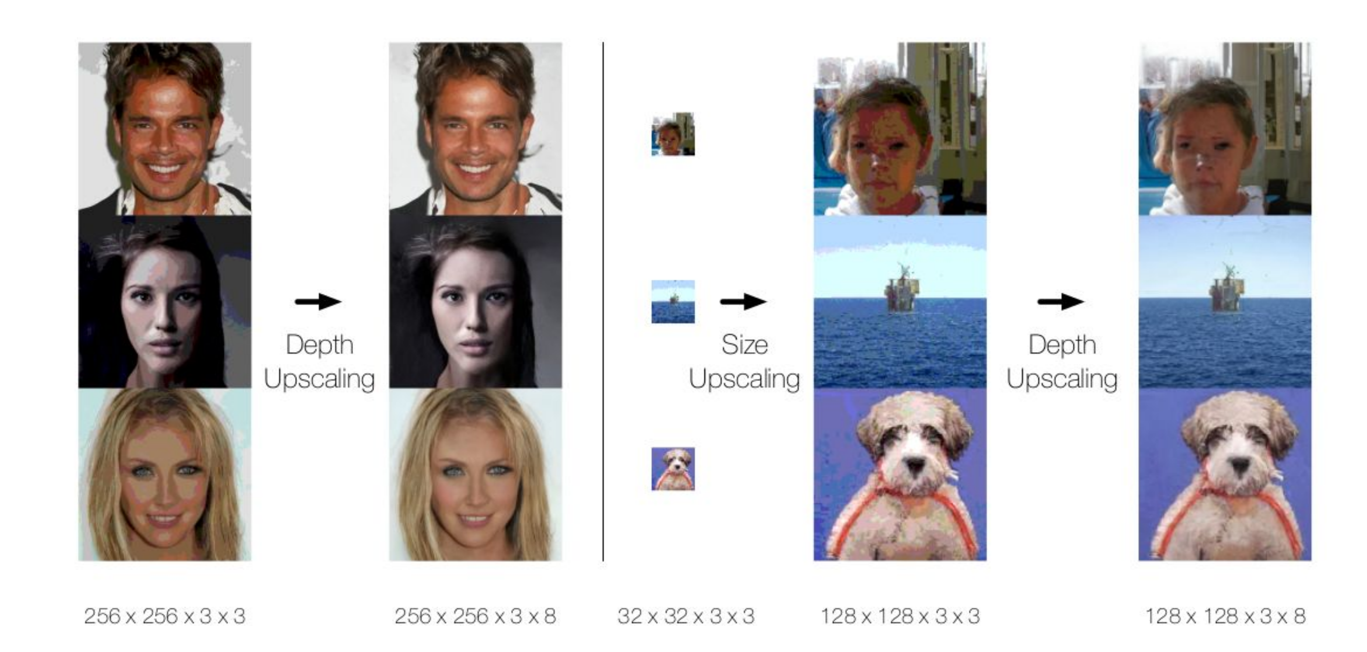
Figure 1. Image completions sampled from a PixelRNN.

Difficulties:

- Order
- Speed

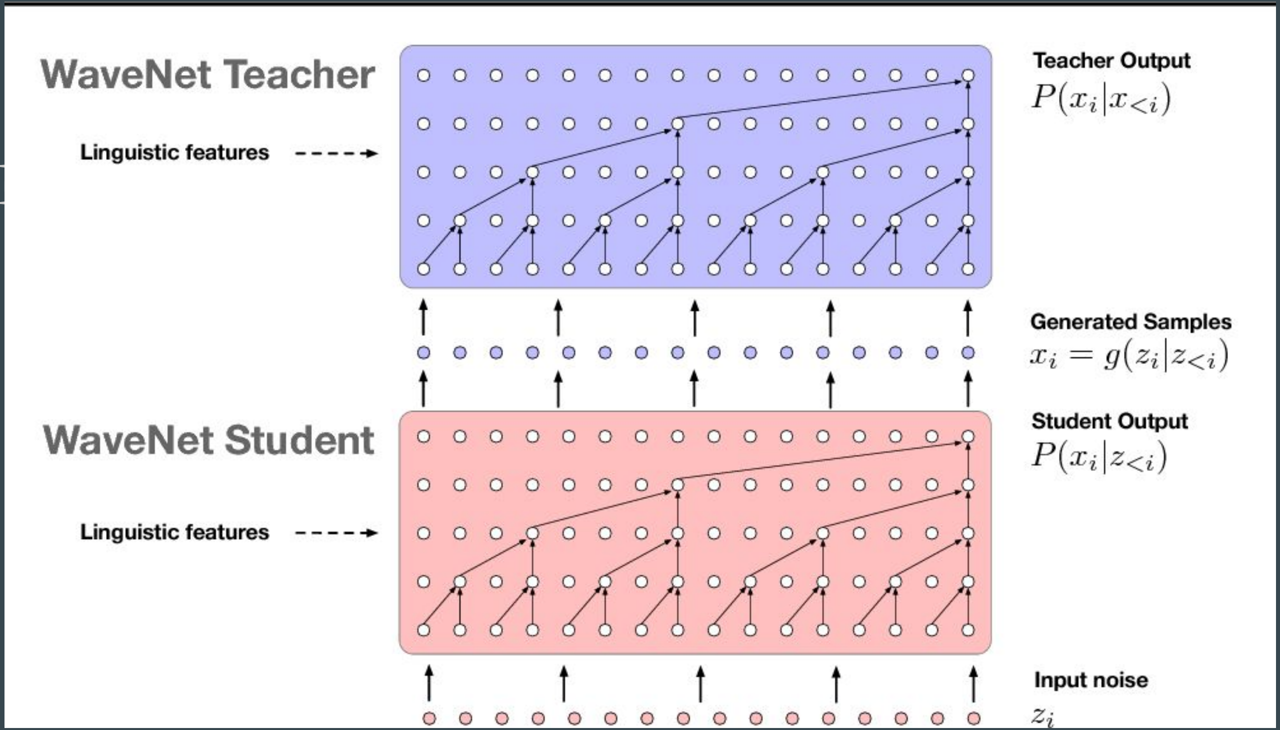
Autoregressive Models IV (image++)

- Subscale pixel networks[6]



Autoregressive Models V (audio)

- Sample RNN[7]
- WaveNet[8]
- Parallel WaveNet[9]



Variational Models I

$$\log P(X) = \log P(X|z) + \log P(z) - \log P(z|X)$$

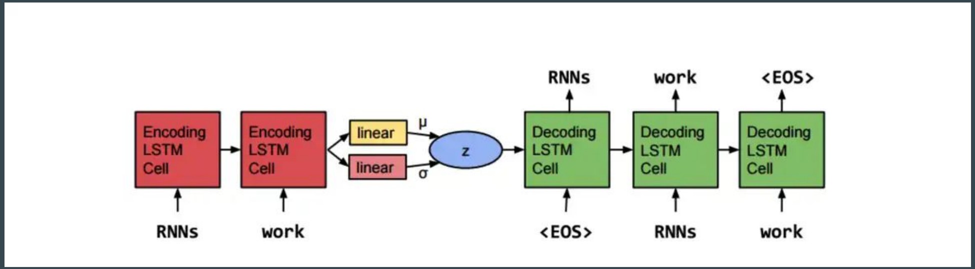
$$E_{z\sim P(z|X)}[\log P(X)] = \log P(X) = E_{z\sim P(z|X)}[\log P(X|z) + \log P(z) - \log P(z|X)]$$

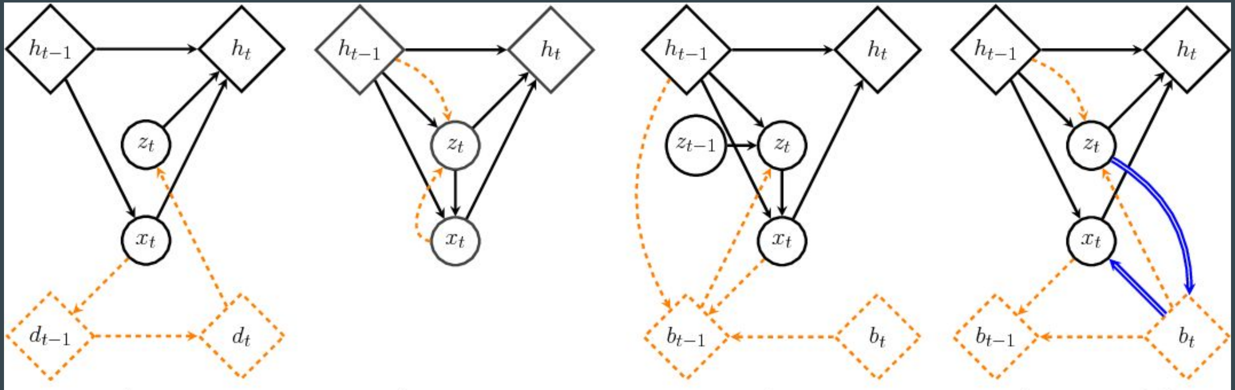
$$\log P(X) = E_{z\sim P(z|X)}[\log P(X|z)] - K_{DL}(P(z|X)||P(z))$$

$$\log P(X) - K_{DL}(Q(z|X)||P(z|X)) = E_{z\sim Q(z|X)}[\log P(X|z)] - K_{DL}(Q(z|X)||P(z))$$

Variational Models II (text)

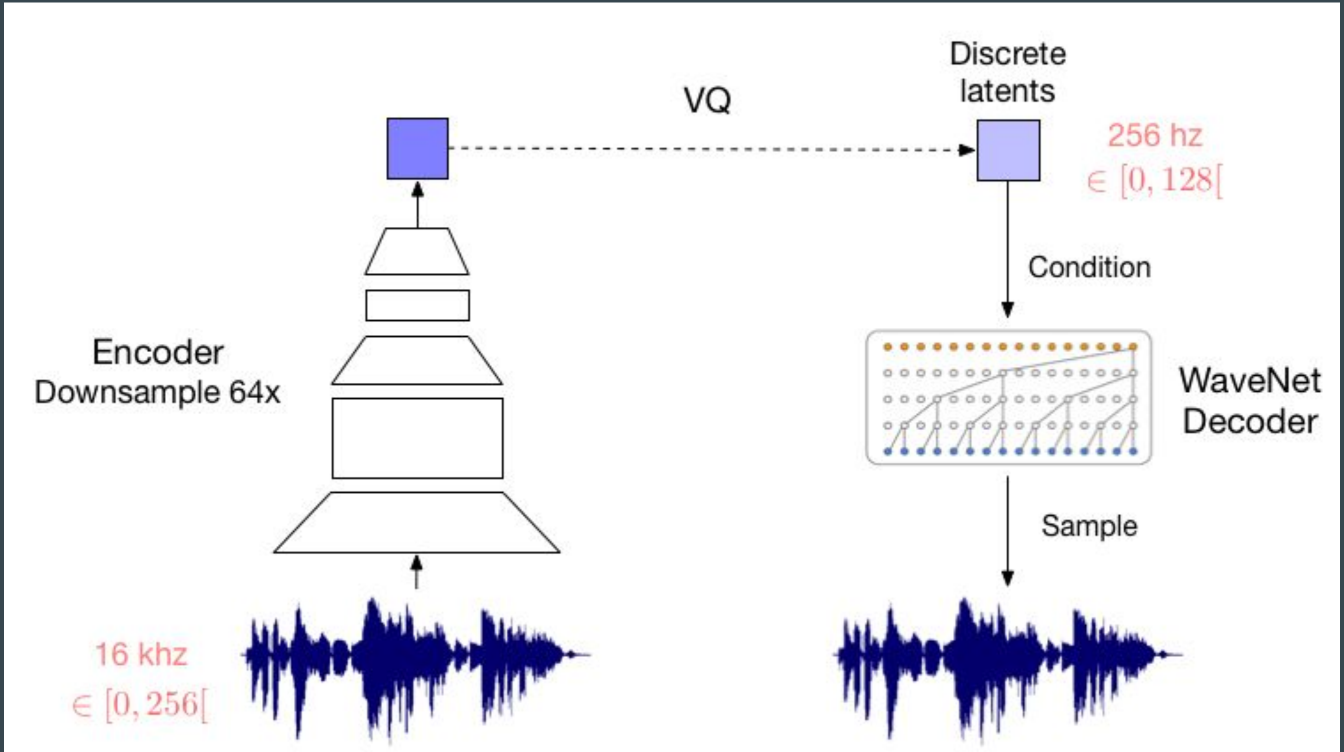
- Text VAE [10]
- Z-Forcing [11]



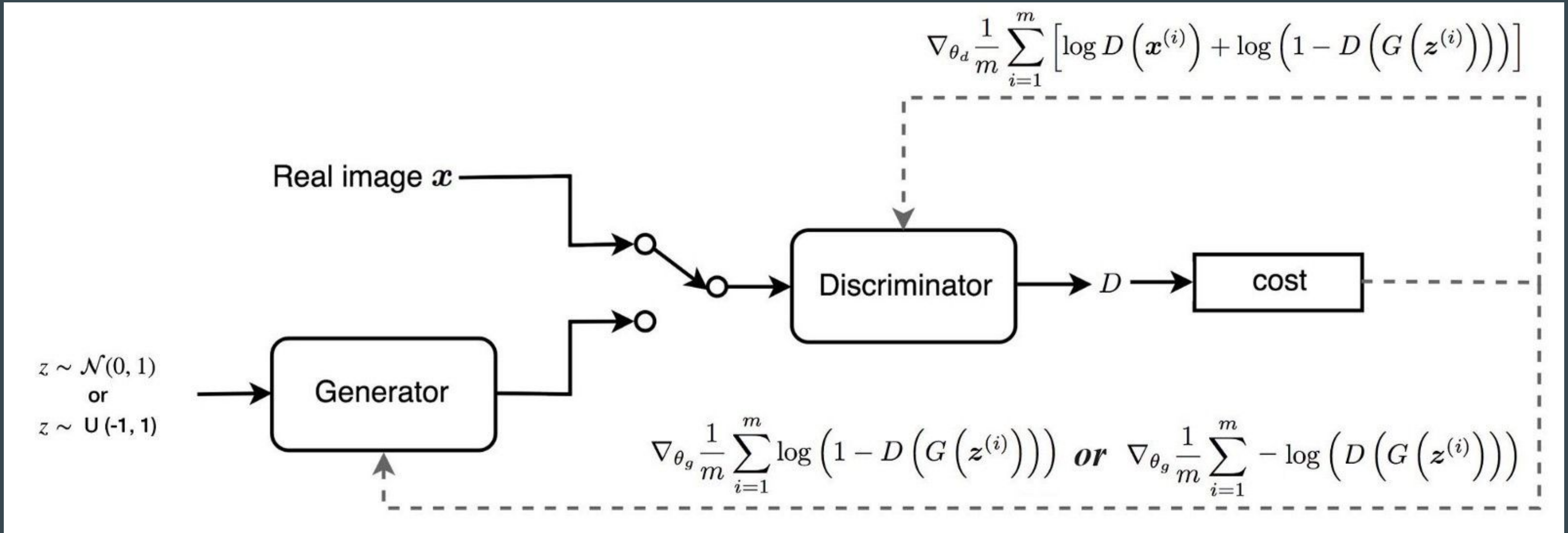


(a) STORN (b) VRNN (c) SRNN (d) Our model

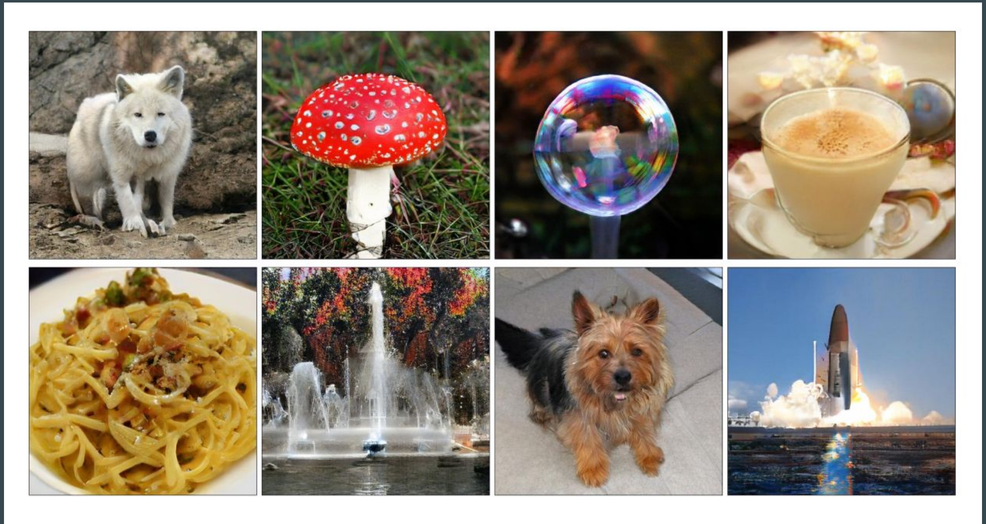
Variational Models III (audio) [12]



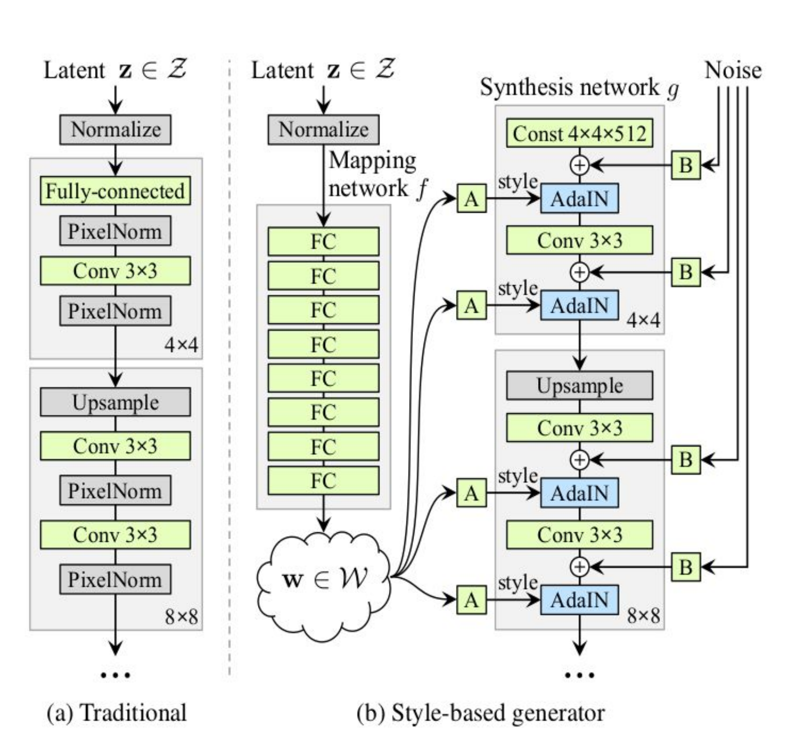
GAN I[13]


$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(\boldsymbol{x}^{(i)}\right) + \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right)\right]$
Real image $\boldsymbol{x}$
Discriminator
D
cost
$z \sim \mathcal{N}(0, 1)$
or
$z \sim$ U (-1, 1)
Generator
$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right)$ ***or*** $\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} -\log\left(D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right)$

GAN II (big-gan)[14]

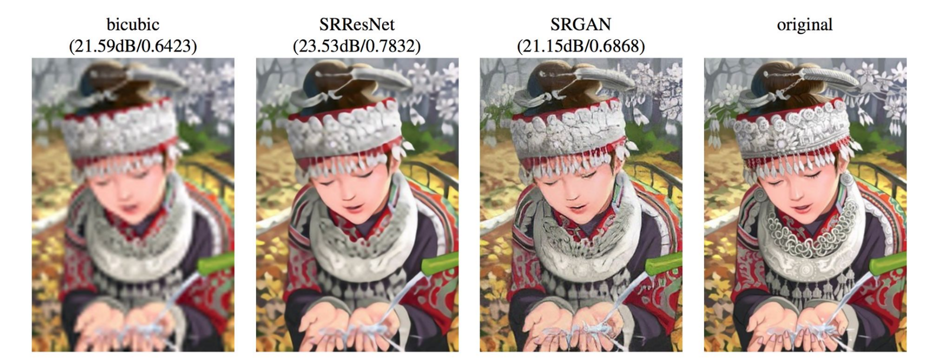


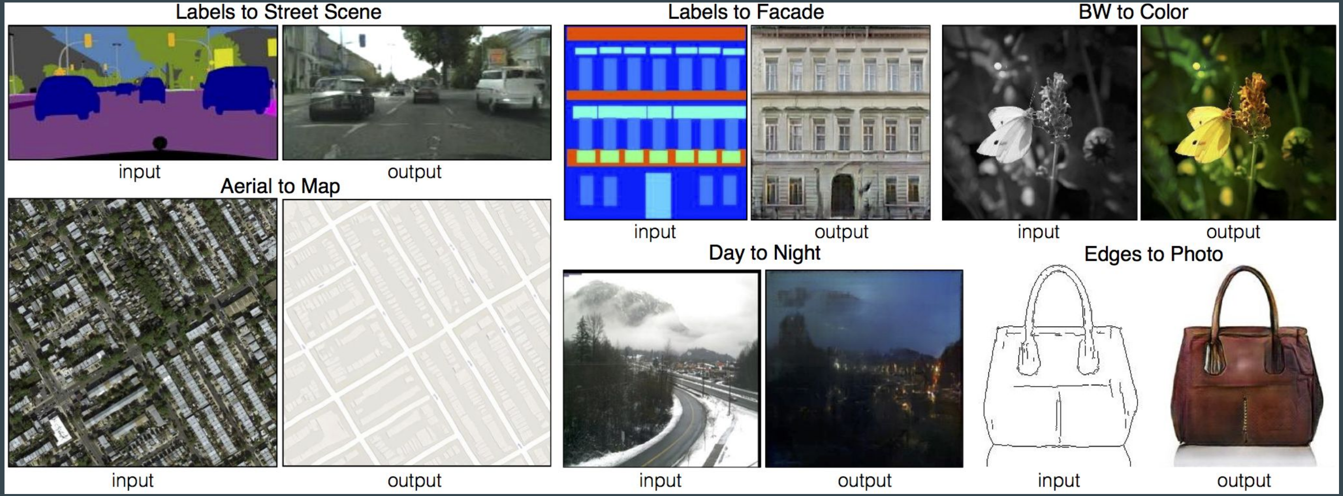
GAN III (pggan++)[15]



GAN IV(use cases)

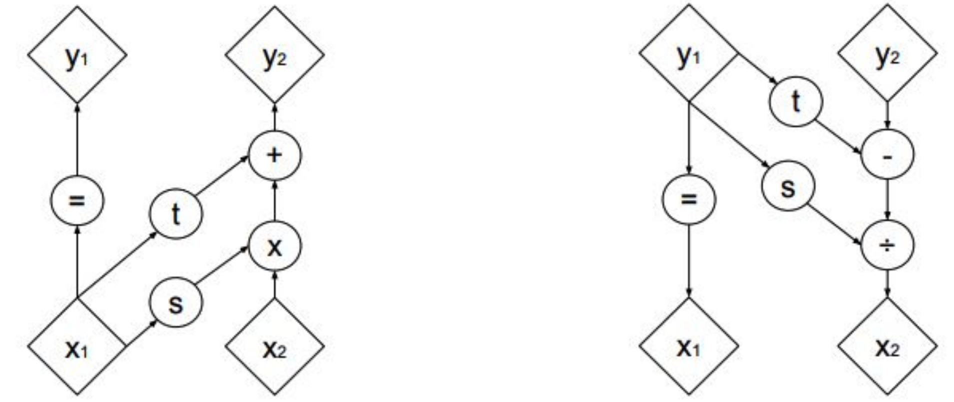
- Pix2pix[16]
- Super res[17]
- Time series prediction[??]





Flow Based Models I

- Real NVP[18]
- Glow[19]
- FloWaveNet[20]



(a) Forward propagation (b) Inverse propagation

Flow Based Models II(ODE)[21, 22]

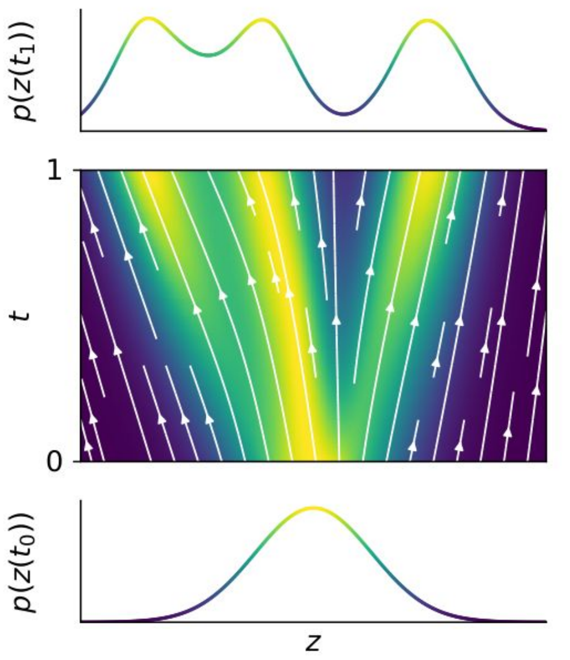


Figure 1: FFJORD transforms a simple base distribution at t_0 into the target distribution at t_1 by integrating over learned continuous dynamics.

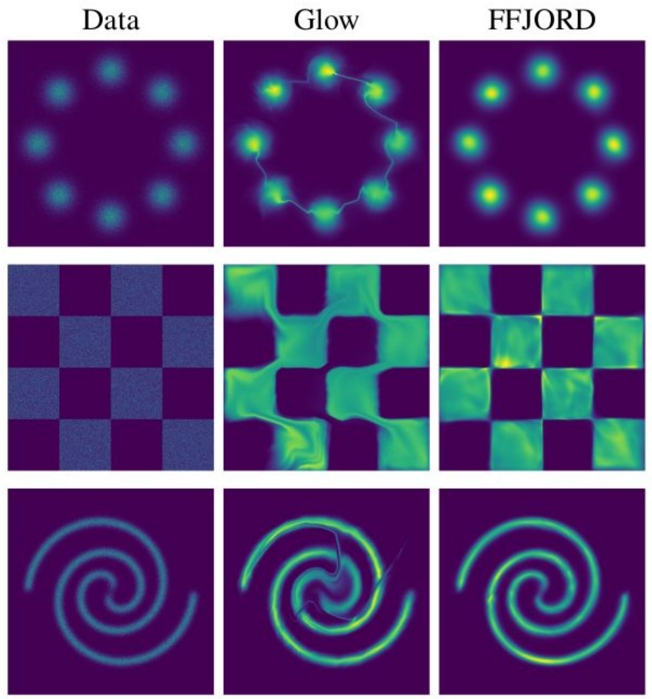
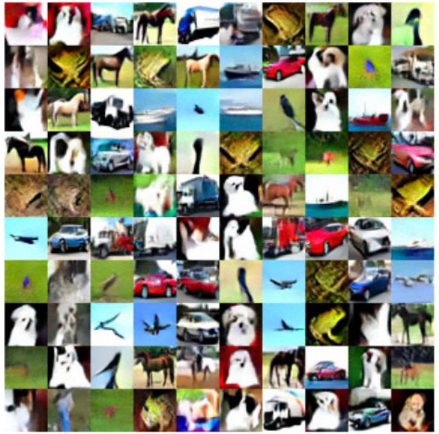


Figure 2: Comparison of trained FFJORD and Glow models on 2-dimensional distributions including multi-modal and discontinuous densities.

Energy Based Models[23]

$$p(x) = \frac{e^{-E(x;\theta)}}{Z(\theta)}; \quad Z(\theta) = \int e^{-E(y;\theta)} dy$$



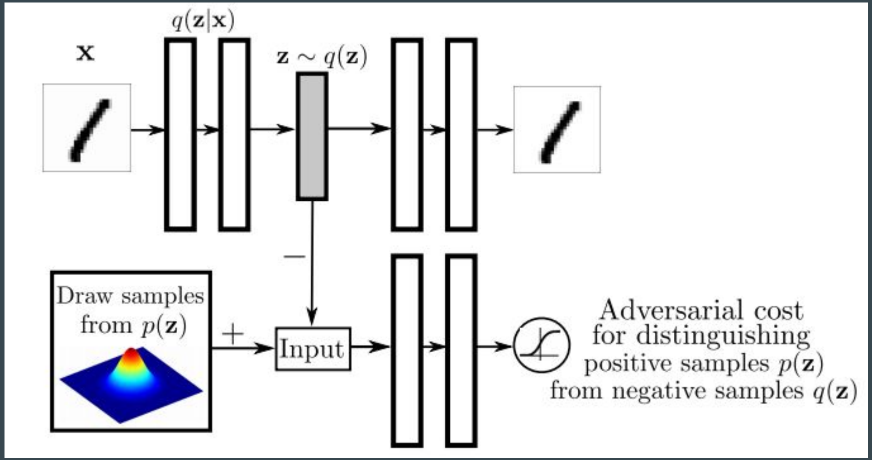
(a) conditional CIFAR10 EBM samples

Model	Inception Score
Unconditional CIFAR10	
PixelCNN	4.60
DCGAN	6.40
Ours (single)	6.43
Ours (10 historical ensemble)	**6.79**
Conditional CIFAR10	
Improved GAN	8.09
Ours	8.52
Spectral Normalization GAN	**8.59**

(b) Table of Inception Scores

Mix & Match

- Adversarial AE[24]
- AliGAN[25]
- VAE + PixelCNN[26]
- ODE + VAE[22]
- EBGAN[27]
- ...



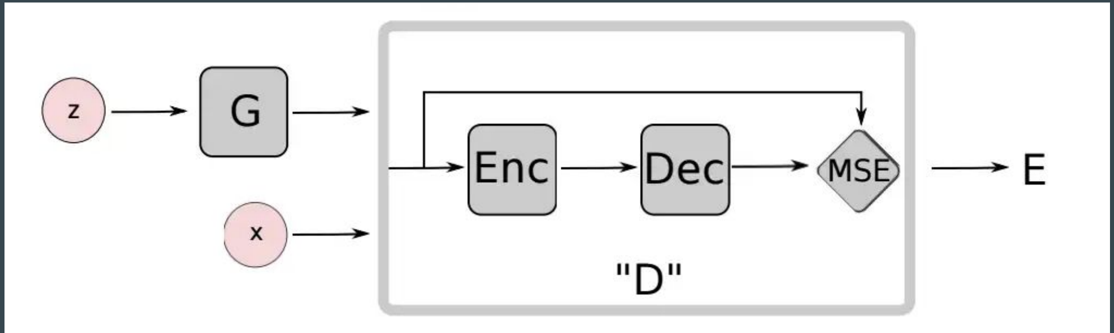


Figure 1: EBGAN architecture.

References I

1. R Adams “A Tutorial on Deep Probabilistic Generative Models”
2. A Ng “On Discriminative vs. Generative classifiers: A comparison of logistic regression and naïve Bayes”
3. Y Kim “Character-Aware Neural Language Models”
4. A Oord “Pixel Recurrent Neural Networks”
5. A Oord “Conditional Image Generation with PixelCNN Decoders”
6. J Menick “Generating High Fidelity Images with Subscale Pixel Networks and Multidimensional Upscaling”
7. S Mehri “SampleRNN: An Unconditional End-to-End Neural Audio Generation Model”

References II

8. A Oord “WaveNet: A Generative Model for Raw Audio”
9. A Oord “Parallel WaveNet: Fast High-Fidelity Speech Synthesis”
10. S Bowman “Generating Sentences from a Continuous Space”
11. A Goyal “Z-Forcing: Training Stochastic Recurrent Networks”
12. A Oord “Neural Discrete Representation Learning”
13. I Goodfellow “Generative Adversarial Networks”
14. A Brock “Large Scale GAN Training for High Fidelity Natural Image Synthesis”

References III

15. T Karras “A Style-Based Generator Architecture for Generative Adversarial Networks”

16. P Isola “Image-to-Image Translation with Conditional Adversarial Networks”

17. X Wang “Image-to-Image Translation with Conditional Adversarial Networks”

18. L Dinh “Density estimation using Real NVP”

19. D Kingma “Glow: Generative Flow with Invertible 1x1 Convolutions”

20. S Kim “FloWaveNet : A Generative Flow for Raw Audio”

21. R Chen “Neural Ordinary Differential Equations”

References IV

22. W Grathwohl “FFJORD: Free-form Continuous Dynamics for Scalable Reversible Generative Models”

23. T Du “Implicit Generation and Generalization with Energy-Based Models”

24. A Makhzani “Adversarial Autoencoders”

25. V Dumoulin “Adversarially Learned Inference”

26. I Gulrajani “PixelVAE: A Latent Variable Model for Natural Images”

27. J Zhao “Energy-based Generative Adversarial Network”