VSB TECHNICAL UNIVERSITY OF OSTRAVA | FACULTY OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE | DEPARTMENT OF COMPUTER SCIENCE

Deep Learning

Generative models

Jan Platoš, Radek Svoboda

March 24, 2024

Department of Computer Science
Faculty of Electrical Engineering and Computer Science
VŠB - Technical University of Ostrava

Generative models

Generative models

Informally:

- **Generative models** can generate new data instances.
- **Discriminative models** discriminate between different kinds of data instances.

Formally:

- **Generative models** capture the joint probability $p(X, Y)$, or just $p(X)$ if there are no labels.
- **Discriminative models** capture the conditional probability $p(Y|X)$.

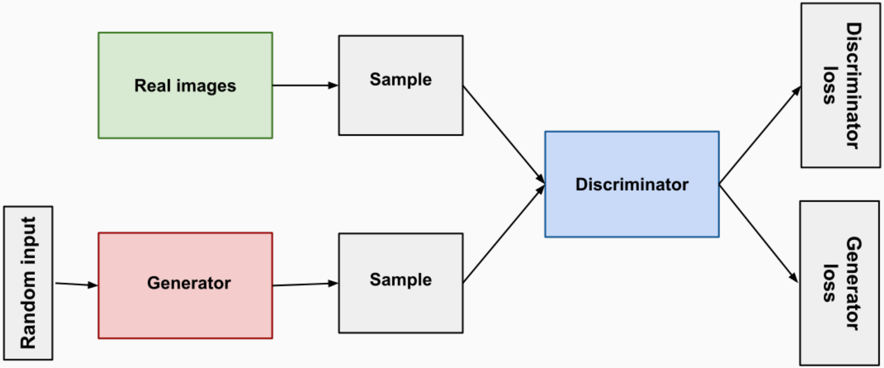
Source: `https://developers.google.com/machine-learning/gan/generative`

Generative models

- Generative models task is much harder than the discriminative.
- Discriminative models identifies the important places for decision.
- Generative models identifies all the relations between objects in the true dataset.
- Learning generative models need more information from data and modified learning process to gather required information.

- GAN is composed of two parts.
- The **generator** learns how to generate realistic data.
- The **discriminator** learns to distinguish between fake and real data.
- The generated images are negative examples for the discriminator.
- The discriminator penalizes the generator if it is not able to generate realistic data.

Generative models - Generative Adversarial Network (GAN)



Schema of the GAN

https://developers.google.com/machine-learning/gan/generative

Generative models - Generative Adversarial Network (GAN) - Discriminator

- The Discriminator is a classifier.
- The goal is to distinguish between the real and generated data.
- The structure depend on the classified data (CNN, Dense, ...).
- The training data comes from:
  - Real data - a real world dataset we try to mimic.
  - Fake data - data generated by the generator.

Generative models - Generative Adversarial Network (GAN) - Discriminator

- The loss functions are two - the discriminator's and the generator's.
- The discriminator's loss is used only in discriminator training.
- The discriminator classifies fake and real data.
- The discriminator's loss penalizes discriminator when classifies real data as fake or fake as real.
- The discriminator updates its weights using back-propagation based on the discriminator's loss.

Generative models - Generative Adversarial Network (GAN) - Generator

- The Generator is a networks that generates data that may be classified as real.
- The Generator is tightly connected to the discriminator.
- The input of the generator is a random vector.
- The input is transformed using the generator network and produces the output.
- The discriminator classifies the generated data.
- The generator loss penalizes the generator for failing to fool the discriminator.

- The generator is directly connected to the output.
- The discriminator network accept the generators output and classify it.
- The resulting loss function back-propagates through discriminator first.
- Then it back-propagates through the generator and updates the weight.

Generative models - Generative Adversarial Network (GAN) - Training

- GAN training proceeds in alternating periods:
- 1. The discriminator trains for one or more epochs.
- 2. The generator trains for one or more epochs.
- Repeat steps 1 and 2 to continue to train the generator and discriminator networks.

- The largest problem is the convergence.
- The improving generator leads into failing discriminator.
- Finally, the discriminator has a 50% accuracy.

- Minmax Loss - maximization of:

$$E_x \left[\log \left(D\left(x\right)\right)\right] + E_z \left[\log \left(1 - D\left(G\left(z\right)\right)\right)\right]$$

- $D(x)$ is the discriminator's estimate of the probability that real data instance x is real.
- E_x is the expected value over all real data instances.
- $G(z)$ is the generator's output when given noise z.
- $D(G(z))$ is the discriminator's estimate of the probability that a fake instance is real.
- E_z is the expected value over all random inputs to the generator (in effect, the expected value over all generated fake instances $G(z)$).
- The formula derives from the cross-entropy between the real and generated distributions.

- Vanishing gradient problem - when discriminator is too good, the generator training can fail due to vanishing gradients. An optimal discriminator doesn't provide enough information for the generator to make progress.
- Modified Minmax Loss - maximization of:

$$\log\left(D\left(G\left(z\right)\right)\right)$$

- Wasserstein Loss

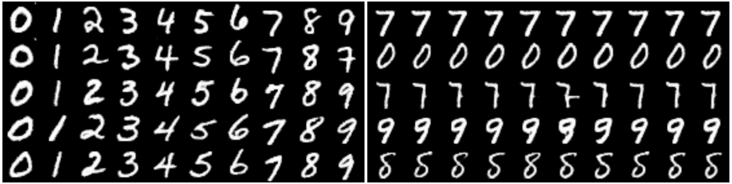
$$D \rightarrow D(x) - D(G(z)) \qquad G \rightarrow D(G(z))$$

Generative models - Generative Adversarial Network (GAN) - Problems

- Mode Collapse - the low variance of the produced images.
- If a generator produces an especially plausible output, the generator may learn to produce only that output.
- In fact, the generator is always trying to find the one output that seems most plausible to the discriminator.
- Each iteration of generator over-optimizes for a particular discriminator, and the discriminator never manages to learn its way out of the trap.
- As a result the generators rotate through a small set of output types. This form of GAN failure is called mode collapse.

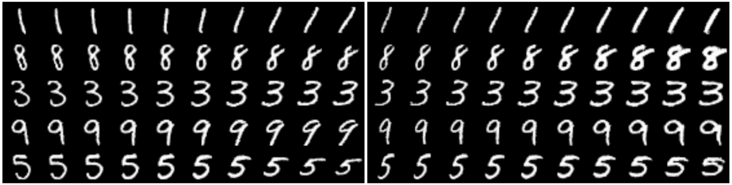
- Deep Convolutional Generative Adversarial Network (DCGAN)
  - An extension of the GAN architecture for using deep convolutional neural networks for both the generator and discriminator models.
- Conditional Generative Adversarial Network (cGAN)
  - An extension to the GAN architecture that makes use of information in addition to the image as input both to the generator and the discriminator models.
  - For example, if class labels are available, they can be used as input.

- Auxiliary Classifier Generative Adversarial Network (AC-GAN)
  - An extension to the GAN that both changes the generator to be class conditional as with the cGAN, and adds an additional or auxiliary model to the discriminator that is trained to reconstruct the class label.
- Information Maximzing Generative Adversarial Network (InfoGAN)
  - An extension to the GAN that attempts to structure the input or latent space for the generator.
  - Specifically, the goal is to add specific semantic meaning to the variables in the latent space.



(a) Varying c_1 on InfoGAN (Digit type)

(b) Varying c_1 on regular GAN (No clear meaning)



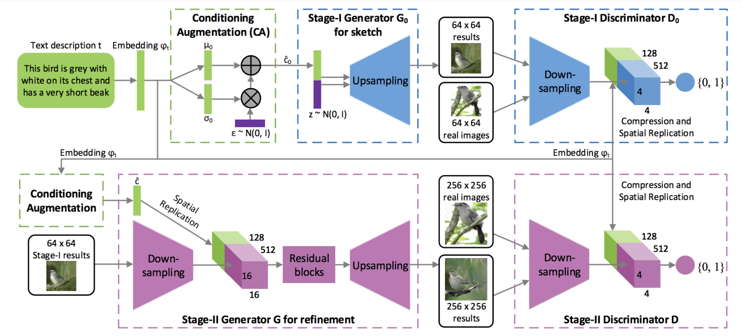
(c) Varying c_2 from -2 to 2 on InfoGAN (Rotation)

(d) Varying c_3 from -2 to 2 on InfoGAN (Width)

Example of the Info GAN

Chen, Xi, et al. "Infogan: Interpretable representation learning by information maximizing generative adversarial nets." Advances in neural information processing systems 29 (2016).

- Stacked Generative Adversarial Network (StackGAN)
    - An extension to the GAN to generate images from text using a hierarchical stack of conditional GAN models.
    - The architecture is comprised of a series of text- and image-conditional GAN models.
    - The first level generator (Stage-I GAN) is conditioned on text and generates a low-resolution image.
    - The second level generator (Stage-II GaN) is conditioned both on the text and on the low-resolution image output by the first level and outputs a high-resolution image.



Archtecture of StackGAN

Zhang, H., et al. "StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks. arXiv e-prints." arXiv preprint arXiv:1612.03242 10 (2016).

- https://machinelearningmastery.com/tour-of-generative-adversarial-network-models
- https://www.youtube.com/watch?v=c-NJtV9Jvp0
- https://openai.com/blog/dall-e/

Questions?