Introduction to Deep Learning (CS474)

Lecture 28

Outline

Module 4

- **Generative Adversarial Networks (GANs)**

Algorithm

Algorithm 1 Minibatch stochastic gradient descent training of generative adversarial nets. The number of steps to apply to the discriminator, k, is a hyperparameter. We used $k = 1$, the least expensive option, in our experiments.

for number of training iterations **do**

    for k steps **do**

        • Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.

        • Sample minibatch of m examples $\{x^{(1)}, \dots, x^{(m)}\}$ from data generating distribution $p_{\text{data}}(x)$.

        • Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(x^{(i)}\right) + \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right) \right].$$

end for

    • Sample minibatch of m noise samples $\{z^{(1)}, \dots, z^{(m)}\}$ from noise prior $p_g(z)$.

    • Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right).$$

end for

The gradient-based updates can use any standard gradient-based learning rule. We used momentum in our experiments.

Deep Convolutional GAN (DCGAN)

DCGAN uses CNN architecture as stable architecture to train GAN. It is achieved by adopting certain architectural constraint to GAN in following ways:

1. Replace pooling layers with strided convolutions in discriminator and transposed convolutions in generator.
2. Use Batch Normalization in both the generator and the discriminator.
3. Remove fully connected layers.
4. Use ReLU activation in generator for all layers except for the output, which uses tanh.
5. Use LeakyReLU activation in the discriminator for all layers.

Slide credit: Dr. G. Krishnamurthy
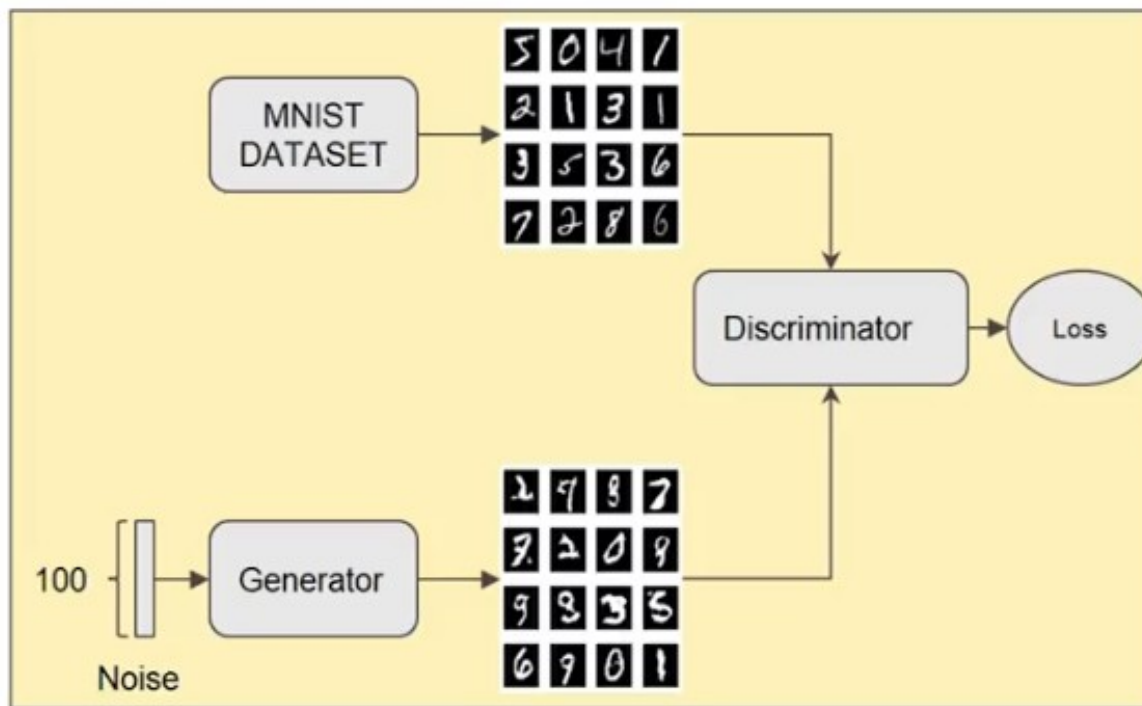
DCGAN Generator Architecture
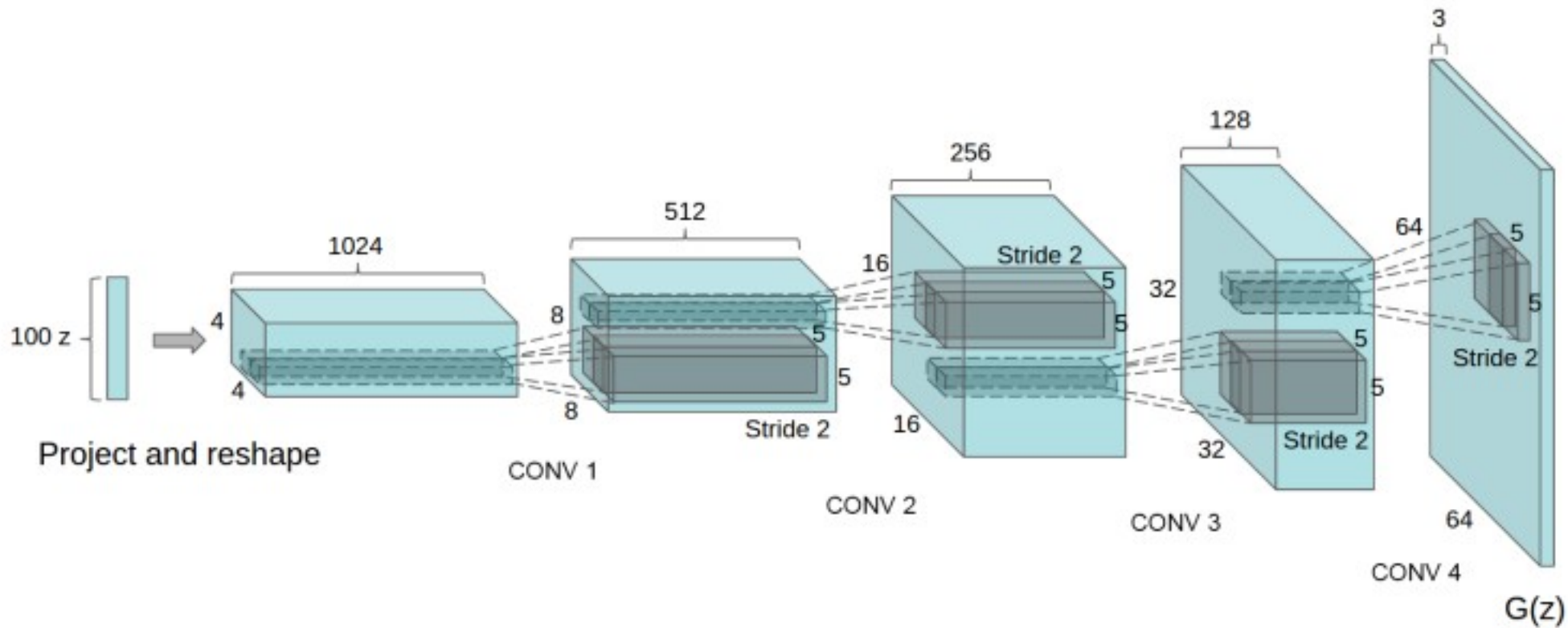
DCGAN Discriminator Architecture

Slide credit: Dr. G. Krishnamurthy
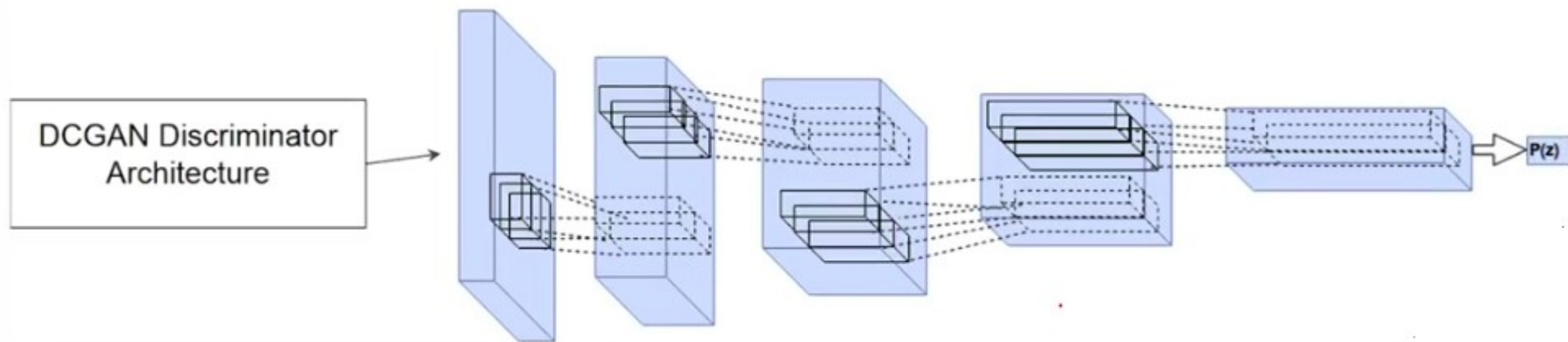
Example

DCGAN with MNIST

Slide credit: Dr. G. Krishnamurthy

References

- All the contents present in the slides are taken from various online resources. Due credit is given in the respective slides.These slides are used for *academic* purposes only.