GENERAL

- Last class. No office hours tomorrow.
- Project Reports Due Sunday Night!
- Keys for PSET 5 can be picked up from Jolley 205 on Tuesday, b/w 12:30pm-1:30pm.

GENERATIVE ADVERSARIAL NETWORKS

Clarification about Loss

- $\max - \log(1 - D(G)) \text{ vs } \min - \log(D(G))$
- Binary Classifier: output of $D$ is $\sigma(y)$ for some $y$
- Discriminator is doing really well, so $D(G)$ is almost 0 ($y << 0$)
- $\max - \log(1 - D(G)) = \min \log(1 - D(G))$
  $$V_y = -(D(G) - 0) = 0 - D(G) \approx 0$$
- $\min - \log(D(G))$
  $$V_y = D(G) - 1 \approx 1$$
- Both are negative (i.e., correct sign, says try to increase $D(G)$)
- But second version has much higher magnitude.
DCGAN: Radford et al.

Generated images from a training set of bedrooms (LSUN dataset).
**GENERATIVE ADVERSARIAL NETWORKS**

Neyshabur et al., Stabilizing GAN Training with Multiple Random Projections

![Diagram of Generative Adversarial Networks](image)

**GENERATIVE ADVERSARIAL NETWORKS**

Neyshabur et al., Stabilizing GAN Training with Multiple Random Projections

![Diagram of Generative Adversarial Networks](image)

**GENERATIVE ADVERSARIAL NETWORKS**

Denton et al., Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks

![Diagram of Generative Adversarial Networks](image)
Figure 1: Our training starts with both the generator (G) and discriminator (D) having a low spatial resolution of $4 \times 4$ pixels. As the training advances, we incrementally add layers to G and D, thus increasing the spatial resolution of the generated images. All existing layers remain trainable throughout the process. Here $[X \times X]$ refers to convolutional layers operating on $X \times X$ spatial resolution. This allows stable synthesis in high resolutions and also speeds up training considerably. On the right we show six example images generated using progressive growing at 1024 x 1024.
UN-SUPERVISED LEARNING

- For a lot of tasks, it is hard to collect enough training data.
- We saw for the stereo example, how you can have an indirect supervision.
- But in other cases, you have to use transfer learning.
  - Train a network on a large dataset for a related task for which you have ground truth.
  - Remove last layer, and use / finetune feature extractor for new task.
- Researchers are exploring tasks made specifically to transfer from.

ADVERSARIAL LOSS

- Train with

\[ L(G) = \|G(x) - y\|^2 - \lambda \log D(G(x)) \]
\[ L(D) = -\log(1 - D(G(x))) - \log D(y) \]

- The GAN loss is unconditional, but there is also a reconstruction loss.
- So the loss says, be close to the true answer, but make your output resemble natural images.
UN-SUPERVISED LEARNING

- Pre-train by learning to add color

Larsson, Maire, Shakhnarovich, CVPR 2017.

UN-SUPERVISED LEARNING

- Pre-train by solving jigsaw puzzles

DOMAIN ADAPTATION

- Generate synthetic training data using renderers.
- But networks trained on synthetic data need not generalize to real data.
- (In fact, they may not transfer from high-quality Flickr data to cell-phone camera data)

Problem Setting

- Have input-output training pairs of \((x', y)\) from source domain: renderings/high-quality images/...
- Have only inputs \(x\) from target domain: where we actually want to use this.
- Train a network so that features computed from \(x'\) and \(x\) have the same distribution ...
  i.e., use GANs!
That's all folks!

- We've covered what forms the foundations of state-of-the-art vision algorithms
  - Will help you read, understand, and implement vision papers
  - But things are changing rapidly: not just new solutions, but new problems
  - So keep reading!

We hope that you have...

- An understanding of the basic math and programming tools to approach vision problems
  - Are as surprised as we are that humans and animals are able to solve this so easily

Reminders

- Fill out course evaluations!
- Repeat Advertisement: CSE 659A in the Spring, "Advances in Computer Vision".