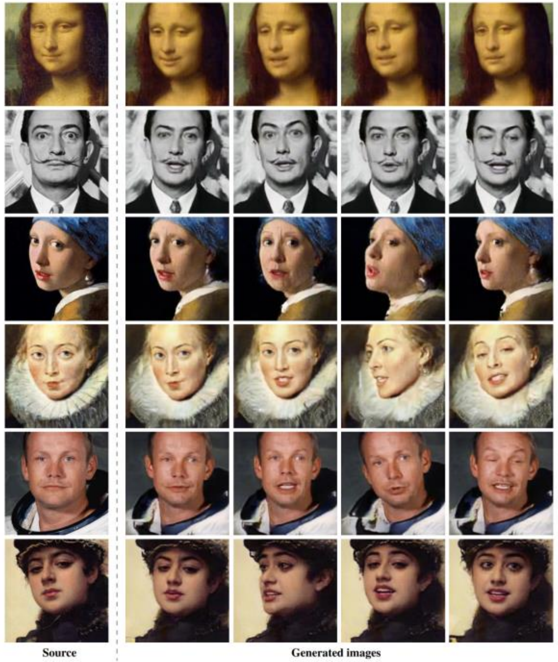
Making GANs Much Better, or If at First You Don't Succeed, Try, Try a GAN

Steve Teig

CEO

Perceive

GANs: Generative Adversarial Networks



GANs: Generative Adversarial Networks

- Generative
  - Produces synthetic output from only a provided, typically random, input

- Network
  - Deep neural network

- Adversarial
  - <u>Should</u> mean “challenging”: examples that are extreme in some way
  - <u>Does</u> mean “produced by a competitive game”: misleading advertising…

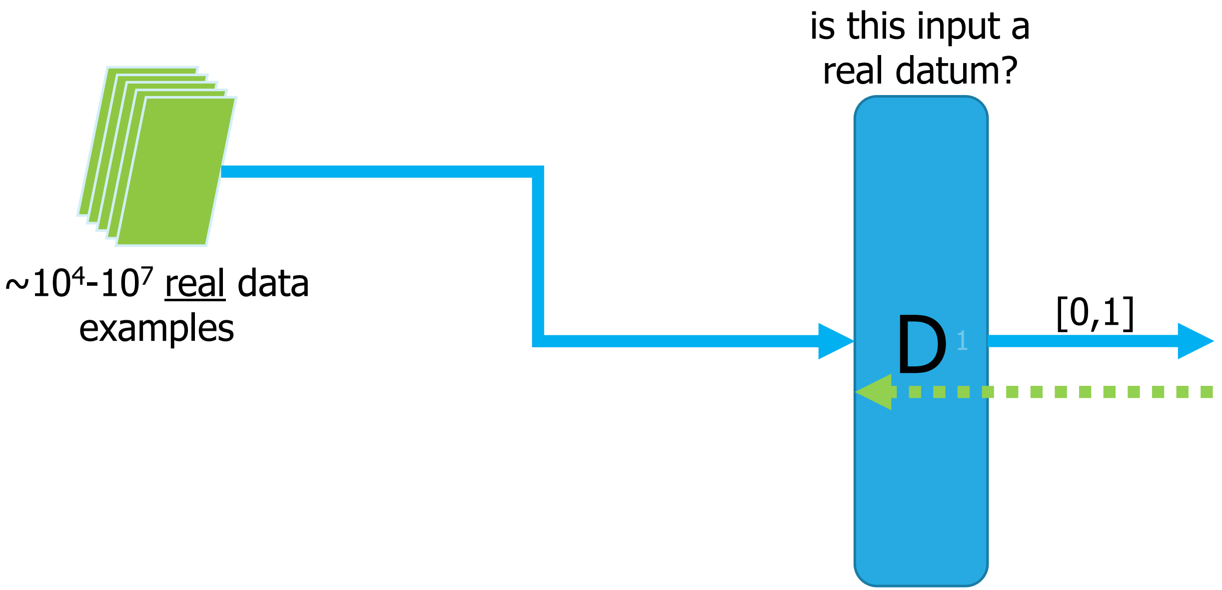
© 2023 Perceive

Why does anyone care about GANs?

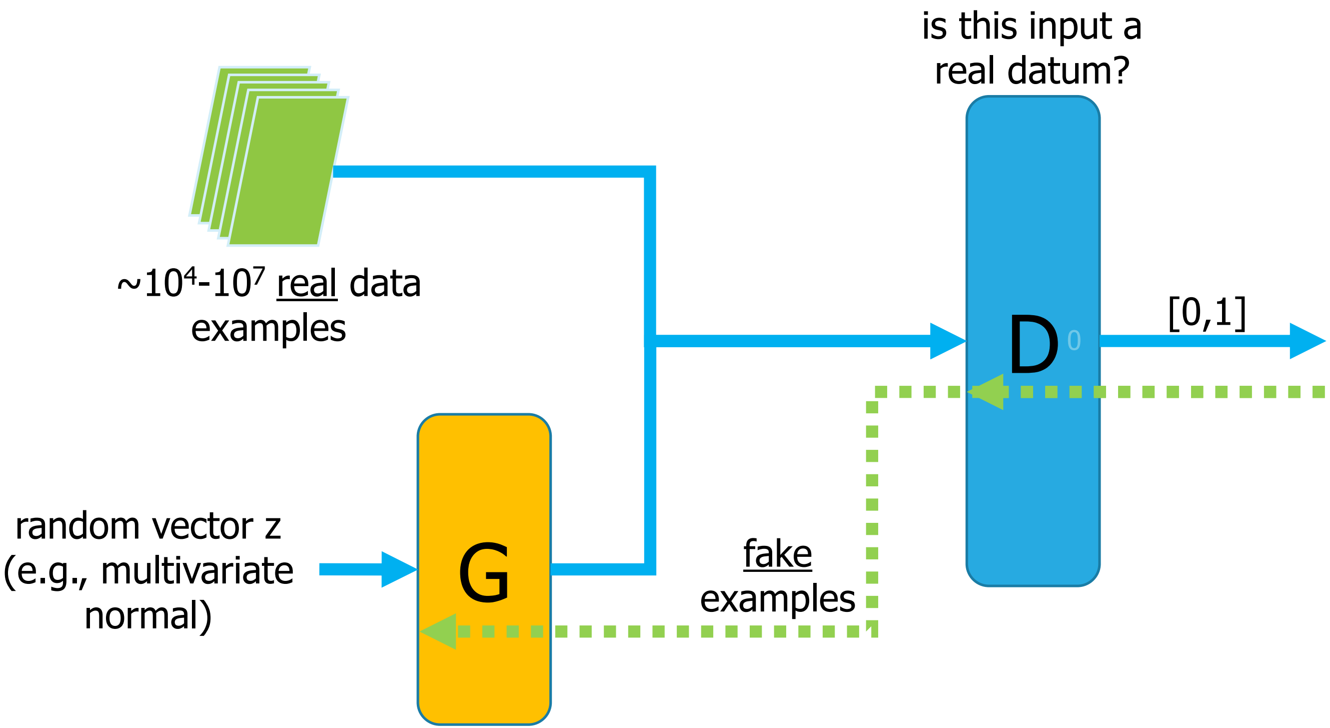
- Data is precious: **no one** has very much of it
  - ~$4x10^{2,500,000}$ possible 4K images
- Can we generate synthetic data that is highly realistic?
  - Enable sophisticated augmentation: avoid overfitting
  - Enable compelling “fakes”: super-resolution, de-noising, art, deepfakes, etc.
- Can we use deep learning to generate realistic, synthetic data?
  - Initially proposed by Goodfellow et al. in 2014; >50K references since then!
  - Try to make synthetic distribution match real distribution

© 2023 Perceive

ABCs of GANs: discriminator network, D



ABCs of GANs: generator network, G

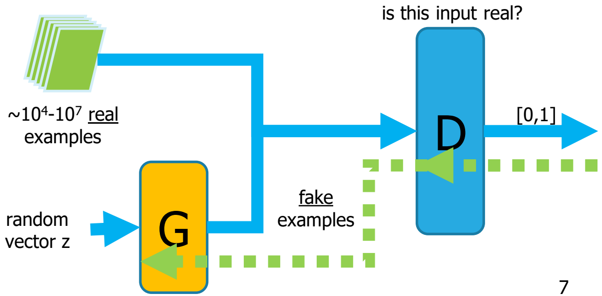


© 2023 Perceive

~~ABCs~~ DEFs of GANs

- GAN "plays a game" where G and D fight
  - Tries to achieve (Nash) equilibrium
- Loss function L
  - $L = \underline{\mathbb{E}_{real\ data\ points\ x}\left[-\log\left(D(x)\right)\right]} + \underline{\mathbb{E}_{random\ vectors\ z}[-\log(1 - D(G(z))]}$
  - Encourage discriminator D to return 1 for real data points
  - Encourage discriminator D to return 0 for fake data points
  - Encourage generator G to learn to fool D; encourage D not to be fooled

- Sounds reasonable, but…
- Underlying reasoning is simplistic → can do <u>much</u> better



© 2023 Perceive

Thinking more critically about GANs

- Why try to match the "distribution" of the labeled input?
  - Copy its biases?!
  - If training set has more "dogs on lawns" than "dogs on streets", should you copy that?
  - If more white than Black people in your training set, should the GAN copy that, too?

- … istribut … n mea …
  - … ave acc … he distr …
  - … number of … tions fr … et could have been …
  - Should every da … y 0?







Perceive

© 2023 Perceive

Thinking more critically about GANs, cont.

- If D seeks a 1 for each training item on average…
  - Then underrepresented features will be ignored to protect the majority!

- If G seeks a 1 for each generated item on average…
  - Then it will focus on "easy" foolers vs. hunting for all foolers

- Is it more important that D get 1's for real data or 0's for fake data?
  - Mainstream GANs treat both as equally important. Why?

A trio of oversights

- GAN loss function is
  - $L = \boxed{\mathbb{E}_{real\ data\ points\ x}\left[-\log\left(D(x)\right)\right]} + \boxed{\mathbb{E}_{random\ vectors\ z}\left[-\log\left(1 - D(G(z))\right)\right]}$

1) First term minimizes D’s <u>average</u> surprise that *real* data are… real
   - Pushes D towards 1 for real data points, but…
   - Small minority of real data points can miss by a mile

2) Second term minimizes D’s <u>average</u> surprise that *fake* data are fake
   - Pushes D towards 0 for fake data points, but…
   - Can get away with “easy” fakes: no driver towards variety

A trio of oversights, cont.

- GAN loss function is
  - $L = \mathbb{E}_{real\ data\ points\ x}\left[-\log\left(D(x)\right)\right] + \mathbb{E}_{random\ vectors\ z}[-log(1 - D(G(z))]$

3) L treats the two terms as <u>equally important</u>
   - Mislabeling a real picture vs. mislabeling a fake picture
   - Better? Worse? Equally bad?
   - Why do you think so?

- ~10^6 real images and ~$10^{2,500,000}$ possible (4K) images → LOTS of ways to generate compelling fakes

© 2023 Perceive

If at first, you don't succeed… try, try a GAN

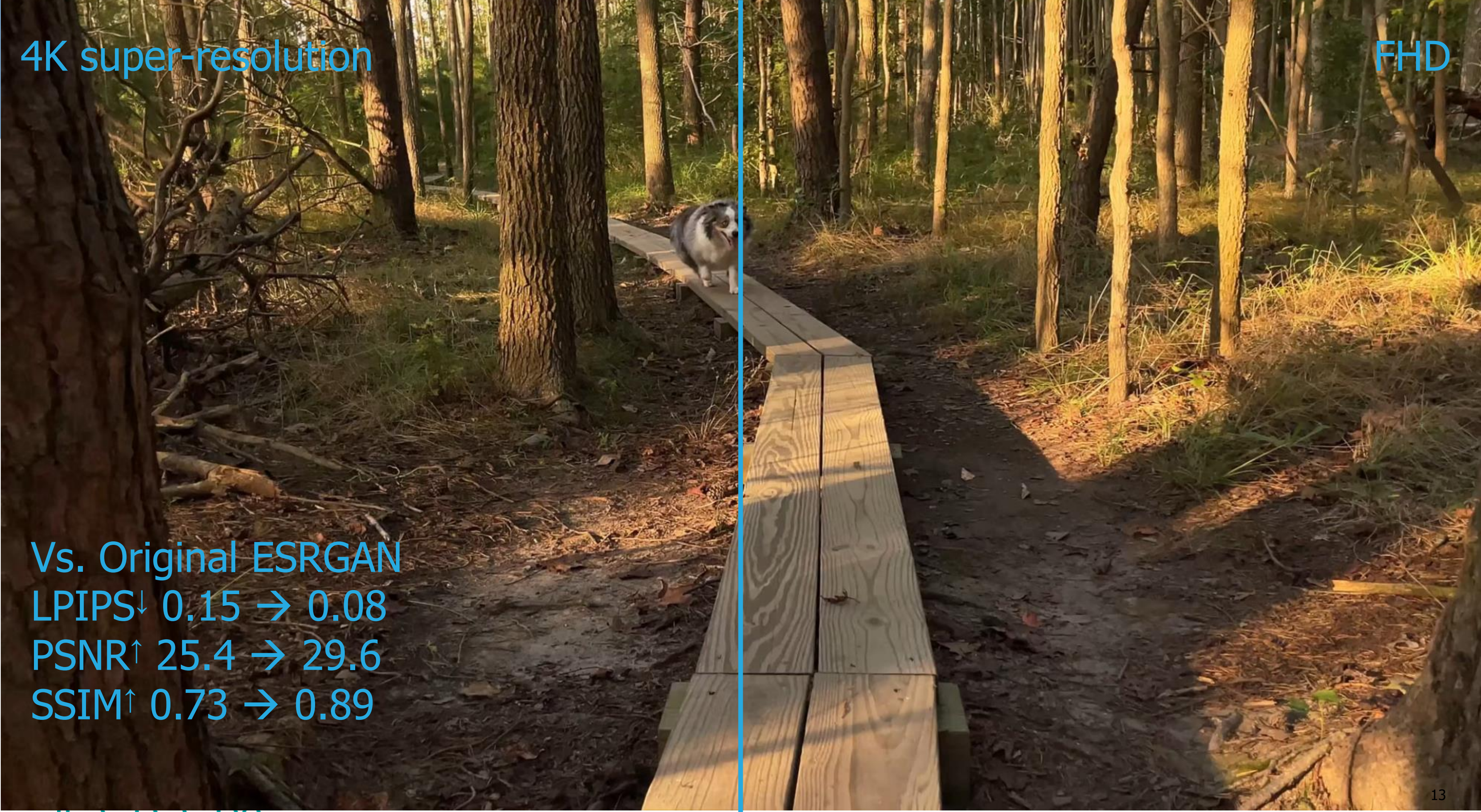
- 1^{st} term: try to make sure D never fails to identify authentic data points
  - $\mathbb{E}_{real\ data\ points\ x}[-log(D(x))] \Rightarrow L_1 = Min_{real\ data\ points\ x}[-log(D(x))]$
  - Continuously differentiable approximation to Min: e.g., LSE

- 2^{nd} term: try to make sure D is never fooled by any synthetic data point
  - I.e., that G searches distribution of z's to get D to make its biggest mistake
  - $\mathbb{E}_{random\ vectors\ z}[-log(1-D(G(z)))] \Rightarrow L_2 = Min_{random\ vectors\ z}[-log(1-D(G(z)))]$

- How should the terms be combined? Can we find an underlying principle?!
  - $L_1 + L_2$? $L_1 + \alpha * L_2$? What's α?
  - $f(L_1,L_2)$? What's f?

Perceive®

© 2023 Perceive


4K super-resolution
FHD
Vs. Original ESRGAN
LPIPS↓ 0.15 → 0.08
PSNR↑ 25.4 → 29.6
SSIM↑ 0.73 → 0.89

Summary

- GANs are based on a powerful insight
  - Synthesis of realistic, fake data by "fooling" training… and people!
- GANs are used widely and are influential
- Unfortunately, mainstream GANs are not "adversarial"
  - Poorly chosen loss function creates severe, unwanted biases
- We can fix that by *minimizing maximum surprise* instead
  - Average accuracy is almost never the right thing for ML
- Used to create state-of-the-art super-resolution at 30 fps with 175x lower power than NVIDIA

© 2023 Perceive

Additional Resources

Resources

Survey of GANs in computer vision:
https://arxiv.org/abs/1906.01529

Original GAN paper:
https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

ESRGAN super-resolution paper:
https://arxiv.org/abs/1809.00219

Perceive: https://www.perceive.io

2023 Embedded Vision Summit

- Women in Vision Reception:
  Tuesday, 6:30-7:30 PM
  Exhibit Floor ET-1
- Perceive exhibit (booth #107)

© 2023 Perceive