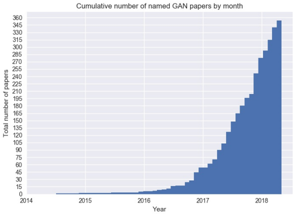
Deep Learning

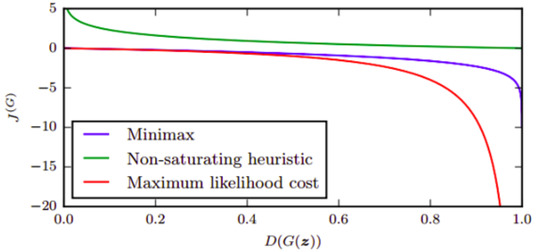
GAN variants



Learning goals

- non-saturating loss
- conditional GANs

NON-SATURATING LOSS

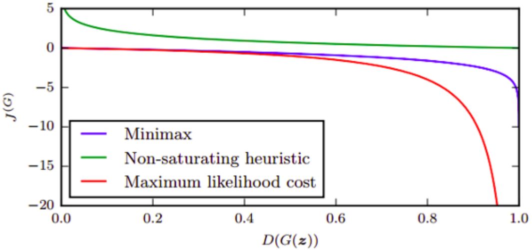


Credit: Daniel Seita

Figure: Various generator loss functions ($J^{(G)}$).

- It was discovered that a relatively strong discriminator could completely dominate the generator.
- When optimizing the minimax loss, as the discriminator gets good at identifying fake images, i.e. as $D(G(\mathbf{z}))$ approaches 0, the gradient with respect to the generator parameters vanishes.

NON-SATURATING LOSS



Credit: Daniel Seita

Figure: Various generator loss functions ($J^{(G)}$).

- Solution: Use a non-saturating generator loss instead: $J^{(G)} = -\frac{1}{2}\mathbb{E}_{\vec{z}\sim p(\vec{z})}[\log D(G(\mathbf{x}))]$
- In contrast to the minimax loss, when the discriminator gets good at identifying fake images, the magnitude of the gradient of $J^{(G)}$ increases and the generator is able to learn to produce better images in successive iterations.

OTHER LOSS FUNCTIONS

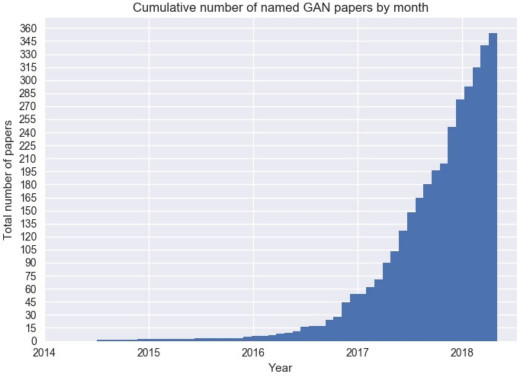
Various losses for GAN training with different properties have been proposed:

GAN	Discriminator Loss	Generator Loss
MM GAN	$\mathcal{L}_{\mathrm{D}}^{\mathrm{GAN}} = -\mathbb{E}_{x\sim p_d}[\log(D(x))] - \mathbb{E}_{\hat{x}\sim p_g}[\log(1 - D(\hat{x}))]$	$\mathcal{L}_{\mathrm{G}}^{\mathrm{GAN}} = \mathbb{E}_{\hat{x}\sim p_g}[\log(1 - D(\hat{x}))]$
NS GAN	$\mathcal{L}_{\mathrm{D}}^{\mathrm{NSGAN}} = -\mathbb{E}_{x\sim p_d}[\log(D(x))] - \mathbb{E}_{\hat{x}\sim p_g}[\log(1 - D(\hat{x}))]$	$\mathcal{L}_{\mathrm{G}}^{\mathrm{NSGAN}} = -\mathbb{E}_{\hat{x}\sim p_g}[\log(D(\hat{x}))]$
WGAN	$\mathcal{L}_{\mathrm{D}}^{\mathrm{WGAN}} = -\mathbb{E}_{x\sim p_d}[D(x)] + \mathbb{E}_{\hat{x}\sim p_g}[D(\hat{x})]$	$\mathcal{L}_{\mathrm{G}}^{\mathrm{WGAN}} = -\mathbb{E}_{\hat{x}\sim p_g}[D(\hat{x})]$
WGAN GP	$\mathcal{L}_{\mathrm{D}}^{\mathrm{WGANGP}} = \mathcal{L}_{\mathrm{D}}^{\mathrm{WGAN}} + \lambda\mathbb{E}_{\hat{x}\sim p_g}[(\|\|\nabla D(\alpha x + (1-\alpha\hat{x})\|\|_2 - 1)^2]$	$\mathcal{L}_{\mathrm{G}}^{\mathrm{WGANGP}} = -\mathbb{E}_{\hat{x}\sim p_g}[D(\hat{x})]$
LS GAN	$\mathcal{L}_{\mathrm{D}}^{\mathrm{LSGAN}} = -\mathbb{E}_{x\sim p_d}[(D(x)-1)^2] + \mathbb{E}_{\hat{x}\sim p_g}[D(\hat{x})^2]$	$\mathcal{L}_{\mathrm{G}}^{\mathrm{LSGAN}} = -\mathbb{E}_{\hat{x}\sim p_g}[(D(\hat{x} - 1)^2]$
DRAGAN	$\mathcal{L}_{\mathrm{D}}^{\mathrm{DRAGAN}} = \mathcal{L}_{\mathrm{D}}^{\mathrm{GAN}} + \lambda\mathbb{E}_{\hat{x}\sim p_d+\mathcal{N}(0,c)}[(\|\|\nabla D(\hat{x})\|\|_2 - 1)^2]$	$\mathcal{L}_{\mathrm{G}}^{\mathrm{DRAGAN}} = \mathbb{E}_{\hat{x}\sim p_g}[\log(1 - D(\hat{x}))]$
BEGAN	$\mathcal{L}_{\mathrm{D}}^{\mathrm{BEGAN}} = \mathbb{E}_{x\sim p_d}[\|\|x - \mathrm{AE}(x)\|\|_1] - k_t\mathbb{E}_{\hat{x}\sim p_g}[\|\|\hat{x} - \mathrm{AE}(\hat{x})\|\|_1]$	$\mathcal{L}_{\mathrm{G}}^{\mathrm{BEGAN}} = \mathbb{E}_{\hat{x}\sim p_g}[\|\|\hat{x} - \mathrm{AE}(\hat{x})\|\|_1]$

Source: Lucic et al. 2016

ARCHITECTURE-VARIANT GANS

Motivated by different challenges in GAN training procedure described, there have been several types of architecture variants proposed. Understanding and improving GAN training is a very active area of research.

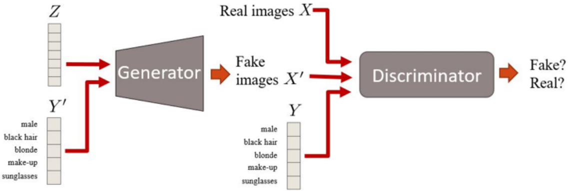


Credit: hindupuravinash

CONDITIONAL GANS: MOTIVATION

- In an ordinary GAN, the only thing that is fed to the generator are the latent variables $\mathbf{z}$.
- A conditional GAN allows you to condition the generative model on additional variables.
- E.g. a generator conditioned on text input (in addition to $\mathbf{z}$) can be trained to generate the image described by the text.

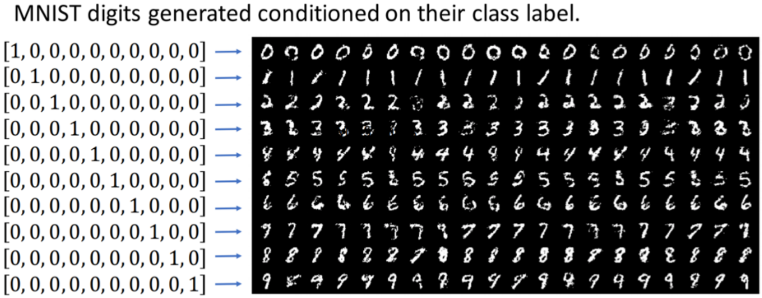
CONDITIONAL GANS: ARCHITECTURE



Credit: Guim Perarnau

- In a conditional GAN, additional information in the form of vector **y** is fed to both the generator and the discriminator.
- **z** can then encode all variations in **z** that are not encoded by **y**.
- E.g. **y** could encode the class of a hand-written number (from 0 to 9). Then, **z** could encode the style of the number (size, weight, rotation, etc).

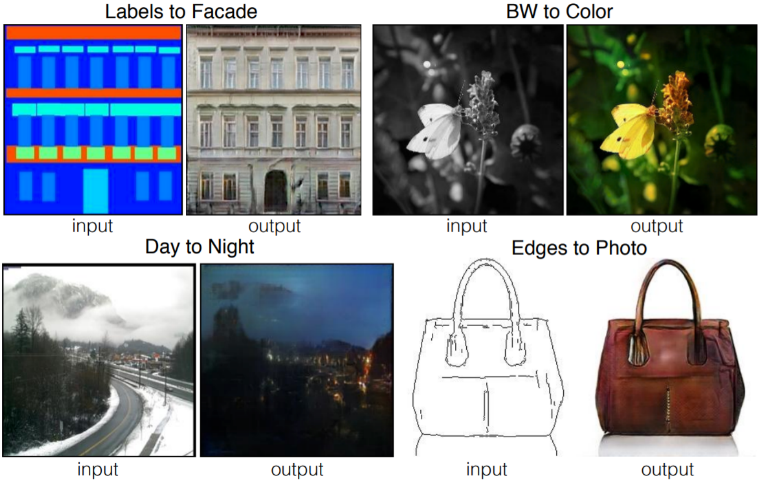
CONDITIONAL GANS: EXAMPLE



Source: Mirza et al. 2014

Figure: When the model is conditioned on a one-hot coded class label, it generates random images that belong (mostly) to that particular class. The randomness here comes from the randomly sampled $\mathbf{z}$. (Note : $\mathbf{z}$ is implicit. It is not shown above.)

CONDITIONAL GANS: MORE EXAMPLES



Source: Isola et al. 2016

Figure: Conditional GANs can translate images of one type to another. In each of the 4 examples above, the image on the left is fed to the network and the image on the right is generated by the network.

MORE GENERATIVE MODELS

- Today, we learned about one kind of (directed) generative models:
- There are other interesting generative models, e.g.:
  - autoregressive models
  - restricted Boltzmann machines.
- Note:
  - It is important to bear in mind that generative models are not a solved problem.
  - There are many interesting hybrid models that combine two or more of these approaches.

REFERENCES

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio (2014)

Generative Adversarial Networks

https://arxiv.org/abs/1406.2661

Santiago Pascual, Antonio Bonafonte, Joan Serra (2017)

SEGAN: Speech Enhancement Generative Adversarial Network

https://arxiv.org/abs/1703.09452

Ian Goodfellow (2016)

NIPS 2016 Tutorial: Generative Adversarial Networks

https://arxiv.org/abs/1701.00160

Lilian Weng (2017)

From GAN to WGAN

https://lilianweng.github.io/lil-log/2017/08/20/from-GAN-to-WGAN.html

REFERENCES

- Mark Chang (2016)
  Generative Adversarial Networks
  https://www.slideshare.net/ckmarkohchang/generative-adversarial-networks

- Lucas Theis, Aaron van den Oord, Matthias Bethge (2016)
  A note on the evaluation of generative models
  https://arxiv.org/abs/1511.01844

- Aiden Nibali (2016)
  The GAN objective, from practice to theory and back again
  https://aiden.nibali.org/blog/2016-12-21-gan-objective/

- Mehdi Mirza, Simon Osindero (2014)
  Conditional Generative Adversarial Nets
  https://arxiv.org/abs/1411.1784

- Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros (2016)
  Image-to-Image Translation with Conditional Adversarial Networks
  https://arxiv.org/abs/1611.07004

REFERENCES

Guim Perarnau (2017)

Fantastic GANs and where to find them

https://guimperarnau.com/blog/2017/03/Fantastic-GANs-and-where-to-find-them