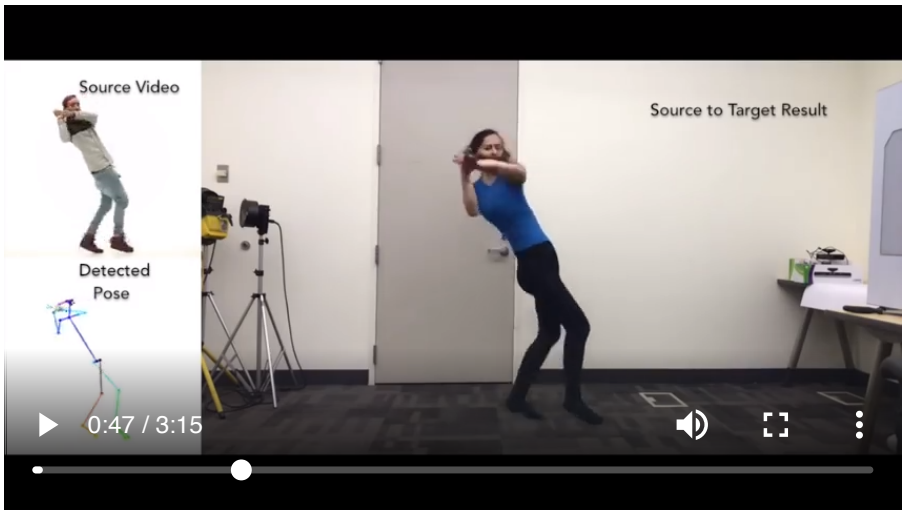
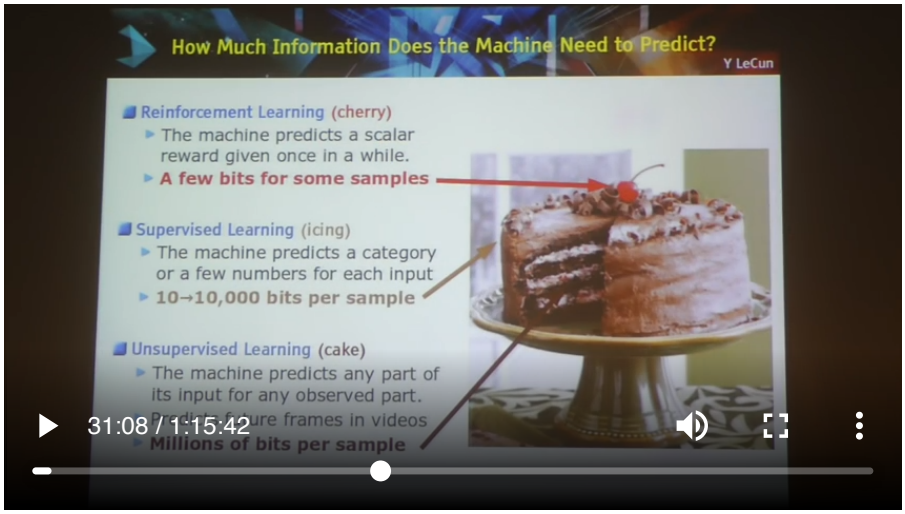
GAN と VAE

- 実習:信長の夢



- Yann LeCun (現 FAIR 所長) said in CMU RI Seminar: Yann LeCun: The Next Frontier in AI: Unsupervised Learning, November 18, 2016

42分20秒頃から **敵対学習は本当にクールなアイデアだ。機械学習の分野で20年来で一番クールだ** 'An adversarial training is a really really really cool idea. It's like the coolest idea in machine learning last twenty years.'



敵対訓練

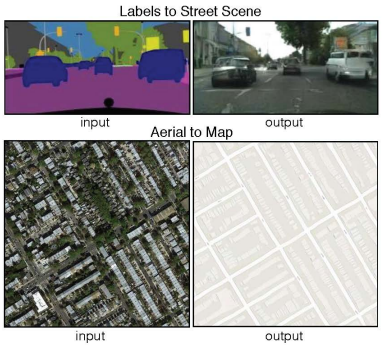
- ざっくりいうと
  - 2人ゼロサムゲーム 古典的なミニマックス戦略
  - プレーヤーは双方ともニューラルネットワーク
- GAN は詳細高密度な画像を生成可能
- GAN はデータの密度関数を学習
- 少数の教師ありデータから，インタラクティブに種々の事例を生成可能 → さまざまな発展の可能性，プライバシー，製品開発，検査，

(Ledig et al 2016)

(Goodfellow 2016)

from Goodfellow(2016)NIPS tutorial

Image to Image Translation



(Isola et al 2016)

(Goodfellow 2016)

from Goodfellow(2016)NIPS tutorial

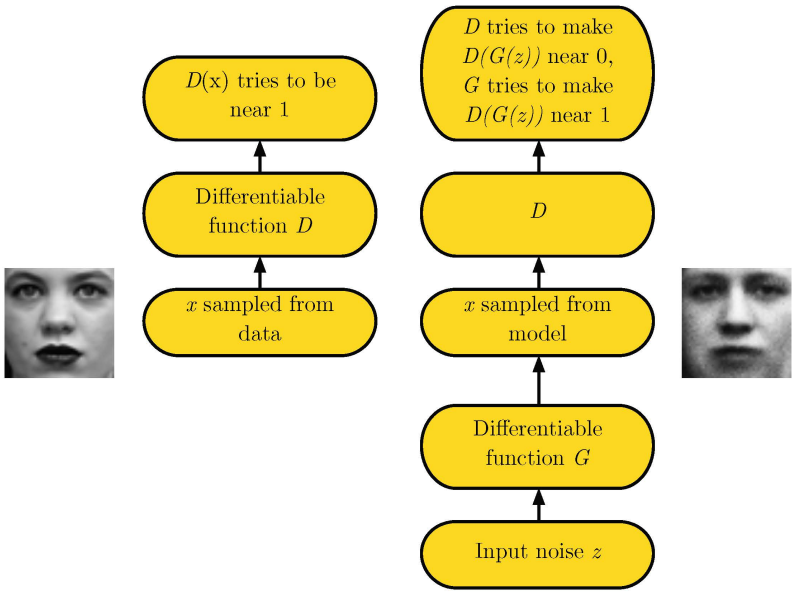
GANs

- Use a latent code
- Asymptotically consistent (unlike variational methods)
- No Markov chains needed
- Often regarded as producing the best samples
  - No good way to quantify this

(Goodfellow 2016)

from Goodfellow(2016)NIPS tutorial

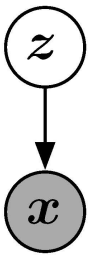
Adversarial Nets Framework



from Goodfellow(2016)NIPS tutorial

Generator Network

$$\boldsymbol{x} = G(\boldsymbol{z}; \boldsymbol{\theta}^{(G)})$$



-Must be differentiable

- No invertibility requirement
- Trainable for any size of ***z***
- Some guarantees require ***z*** **to have higher dimension than** *x*
- **Can make *x* conditionally Gaussian given** *z* **but need not do so**

from

(Goodfellow 2016)

Goodfellow(2016)NIPS tutorial

Minimax Game

$$J^{(D)} = -\frac{1}{2}\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}} \log D(\boldsymbol{x}) - \frac{1}{2}\mathbb{E}_{\boldsymbol{z}} \log\left(1 - D\left(G(\boldsymbol{z})\right)\right)$$
$$J^{(G)} = -J^{(D)}$$

-Equilibrium is a saddle point of the discriminator loss
-Resembles Jensen-Shannon divergence
-Generator minimizes the log-probability of the discriminator being correct

from Goodfellow(2016)NIPS tutorial

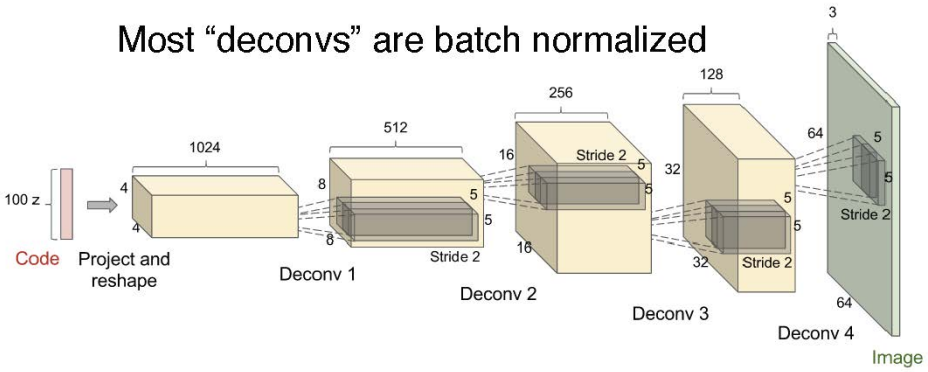
Non-Saturating Game

$$J^{(D)} = -\frac{1}{2}\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}} \log D(\boldsymbol{x}) - \frac{1}{2}\mathbb{E}_{\boldsymbol{z}} \log\left(1 - D\left(G(\boldsymbol{z})\right)\right)$$
$$J^{(G)} = -\frac{1}{2}\mathbb{E}_{\boldsymbol{z}} \log D\left(G(\boldsymbol{z})\right)$$

-Equilibrium no longer describable with a single loss
-Generator maximizes the log-probability of the discriminator being mistaken
-Heuristically motivated; generator can still learn even when discriminator successfully rejects all generator samples

from Goodfellow(2016)NIPS tutorial

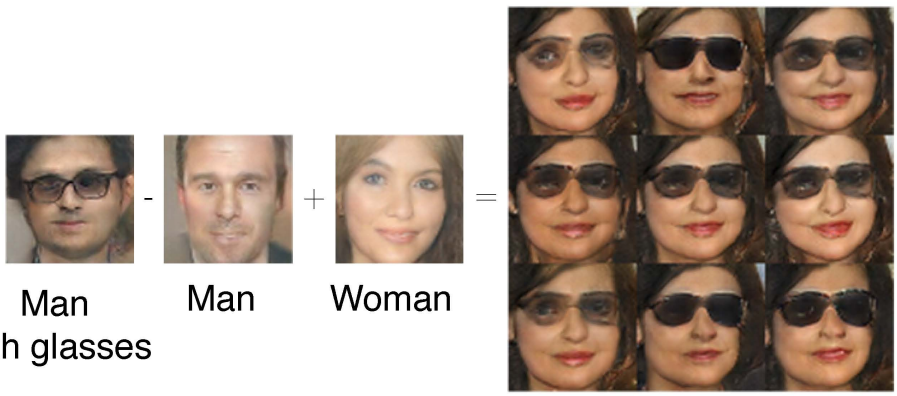
DCGAN Architecture



(Radford et al 2015)

from Goodfellow(2016)NIPS tutorial

Vector Space Arithmetic

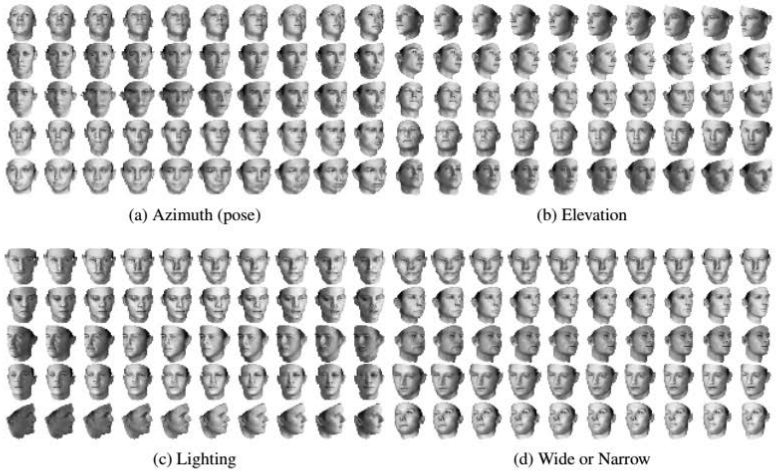


from

(Radford et al, 2015)

Goodfellow(2016)NIPS tutorial

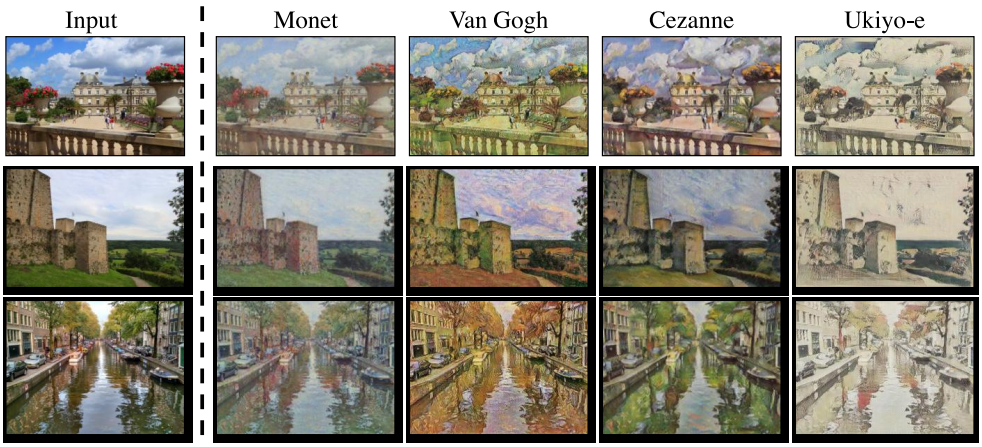
Learning interpretable latent codes / controlling the generation process



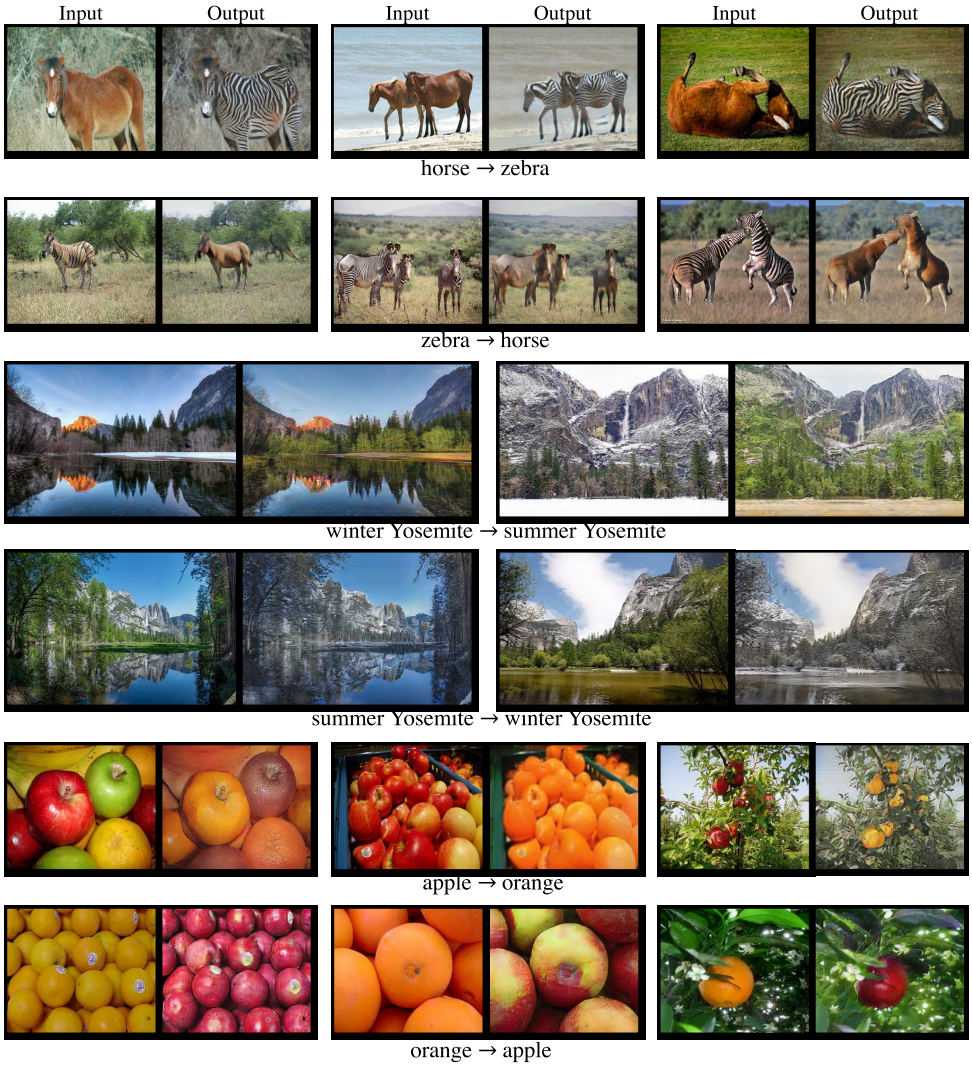
InfoGAN (Chen et al 2016)

(Goodfellow 2016)

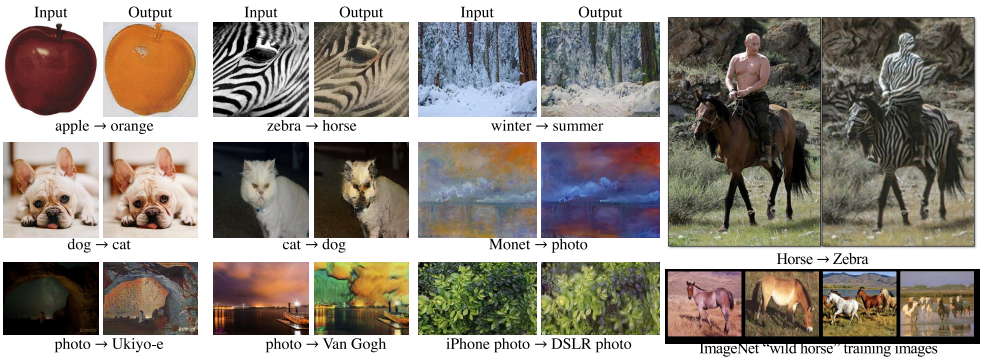
from Goodfellow(2016)NIPS tutorial



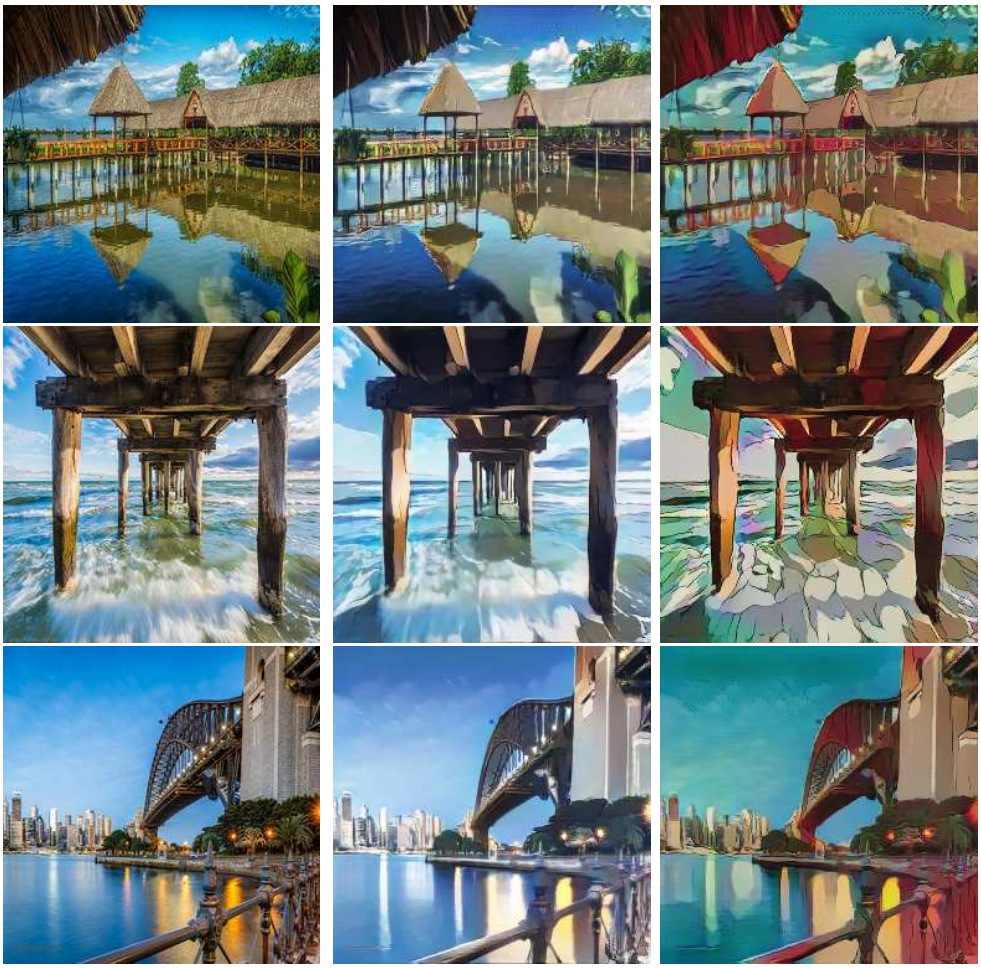
from Zhu et. al.(2017) CycleGAN



from Zhu et. al.(2017) Cycle GAN



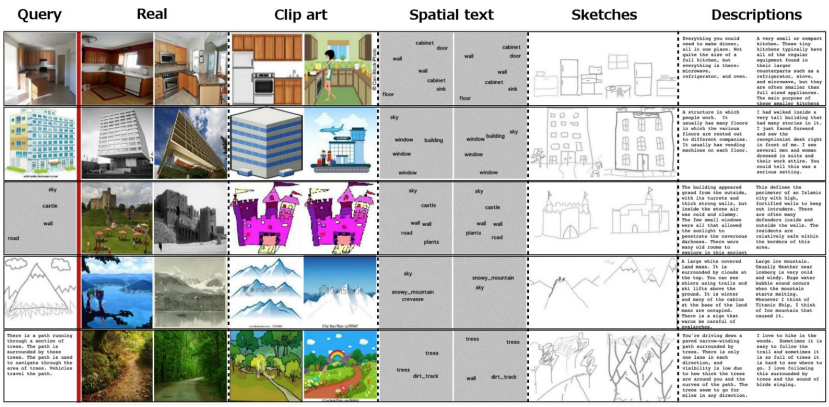
from Zhu et. al.(2017) Cycle GAN



(a) input photo (b) Shinkai style (c) Hayao style

from Chen et. al.(2018) CartoonGAN

Cross-domain retrieval

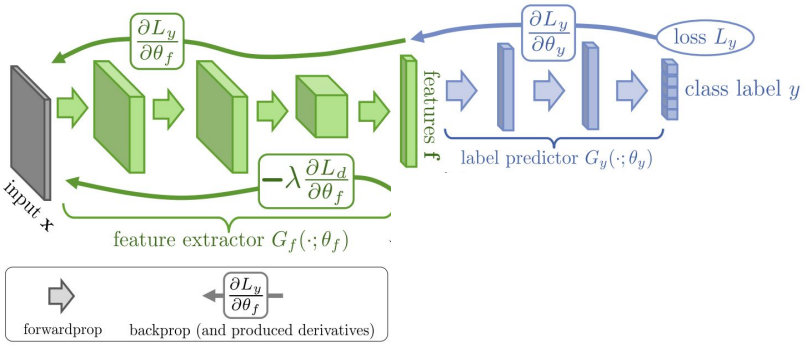


1. Castrejon & Aytar et al. “Learning Aligned Cross-Modal Representations from Weakly Aligned Data”. In CVPR, 2016.

Trends: Domain Adaptation / Unsupervised Alignment

from Reed et.al,. (2017)

Unsupervised domain transfer for classification

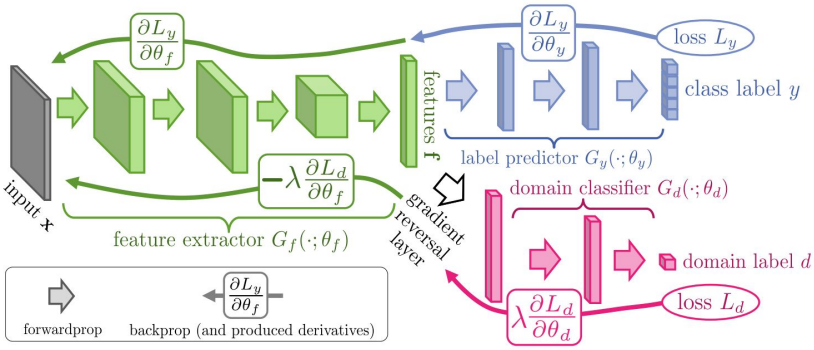


1. *Ganin et al.* "Domain-Adversarial Training of Neural Networks", JMLR 2016

Trends: Domain Adaptation / Unsupervised Alignment

from Reed et.al,. (2017)

Unsupervised domain transfer for classification

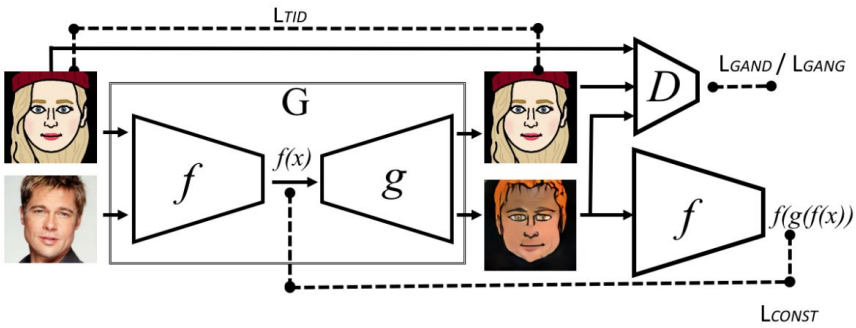


1. *Ganin et al.* "Domain-Adversarial Training of Neural Networks", JMLR 2016

Trends: Domain Adaptation / Unsupervised Alignment

from Reed et.al,. (2017)

Unsupervised cross-domain image generation



1. Taigmen et al. *"Unsupervised Cross-domain image generation".* In *ICLR 2017.*

Trends: Domain Adaptation / Unsupervised Alignment

from Reed et.al,. (2017)

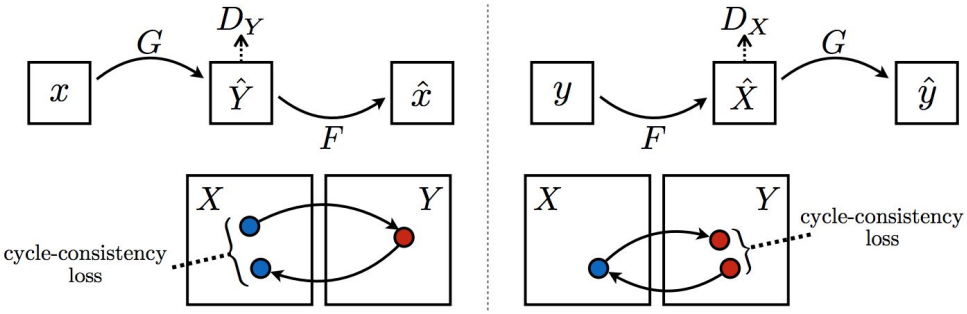
Unsupervised cross-domain image generation

1. Taigmen et al. *"Unsupervised Cross-domain image generation"*. In *ICLR 2017*.

Trends: Domain Adaptation / Unsupervised Alignment

from Reed et.al,. (2017)

Cycle-consistency loss



1. Zhu et al. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks". In ICCV, 2017.

Trends: Domain Adaptation / Unsupervised Alignment

from Reed et.al,. (2017)

- Collection of Interactive Machine Learning Examples

Machine Learning Examples: Seedbank

To see end-to-end examples of the interactive machine learning analyses that Colaboratory makes possible, check out the Seedbank project.

A few featured examples:

- Neural Style Transfer: Use deep learning to transfer style between images.
- EZ NSynth: Synthesize audio with WaveNet auto-encoders.
- Fashion MNIST with Keras and TPUs: Classify fashion-related images with deep learning.
- DeepDream: Produce DeepDream images from your own photos.
- Convolutional VAE: Create a generative model of handwritten digits.
- Convolutional VAE
- deep dream
- deep dream on google drive