Discriminative Deep Networks

“Rockfish”
Discriminative Deep Networks

Raw, Unlabeled Pixels
Generative Deep Networks

Raw, Unlabeled Pixels
Grayscale image: $L$ channel

$X \in \mathbb{R}^{H \times W \times 1}$

Color information: $ab$ channels

$\hat{Y} \in \mathbb{R}^{H \times W \times 2}$

Zhang, Isola, Efros. **Colorful Image Colorization.** In ECCV, 2016.
Grayscale image: $L$ channel

$X \in \mathbb{R}^{H \times W \times 1}$

Concatenate $(L, ab)$ channels

$(X, \hat{Y})$

Simple L2 regression doesn’t work 😞

$$L_2(\hat{Y}, Y) = \frac{1}{2} \sum_{h,w} ||Y_{h,w} - \hat{Y}_{h,w}||^2_2$$
\[ L_2(\hat{Y}, Y) = \frac{1}{2} \sum_{h,w} \| Y_{h,w} - \hat{Y}_{h,w} \|^2 \]
Better Loss Function

\[ \theta^* = \underset{\theta}{\arg \min} \ell(\mathcal{F}_\theta(X), Y) \]

- Regression with L2 loss inadequate
  \[ L_2(\hat{Y}, Y) = \frac{1}{2} \sum_{h,w} \| Y_{h,w} - \hat{Y}_{h,w} \|_2^2 \]
- Use per-pixel multinomial classification
  \[ L(\hat{Z}, Z) = -\frac{1}{HW} \sum_{h,w} \sum_q z_{h,w,q} \log(\hat{Z}_{h,w,q}) \]
Designing loss functions

Input  | Zhang et al. 2016  | Ground truth

Color distribution cross-entropy loss with colorfulness enhancing term.

[Zhang, Isola, Efros, ECCV 2016]
Universal loss?
Generative Adversarial Network (GANs)

Real photos

Generated images

Generated vs Real (classifier)

[Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, Bengio 2014]
Generator

[Goodfellow et al., 2014]
G tries to synthesize fake images that fool D

D tries to identify the fakes

[Goodfellow et al., 2014]
G’s perspective: \( D \) is a loss function.

Rather than being hand-designed, it is \textit{learned}.

[Goodfellow et al., 2014]
[Isola et al., 2017]
BW → Color

Data from [Russakovsky et al. 2015]
BW → Color

Data from [Russakovsky et al. 2015]
Data from [maps.google.com]
Data from [maps.google]
Labels → Facades

Input

Output

Data from [Tylecek, 2013]
Labels → Facades

Data from [Tylecek, 2013]
Day → Night

Data from [Laffont et al., 2014]
Edges $\rightarrow$ Images

Edges from [Xie & Tu, 2015]
Sketches $\rightarrow$ Images

Trained on Edges $\rightarrow$ Images

Data from [Eitz, Hays, Alexa, 2012]
Image-to-image translation in PyTorch (e.g., horse2zebra, edges2cats, and more)

- pytorch
- gan
- cycleran
- pix2pix
- deep-learning
- computer-vision
- computer-graphics
- image-manipulation
- image-generation

- generative-adversarial-network
- gans

- 223 commits
- 3 branches
- 0 releases
- 26 contributors

Branch: master  New pull request

- taesung89 Update README.md
- Latest commit 699e173 10 days ago
- data
- datasets
- imgs
- models

1. datasets are now configured automatically based on dataset_mode op...

Multiple changes regarding option management. See below.

add edges2cats demo

TestModel now supports model_suffix option that can change the name o...
#edges2cats  [Christopher Hesse]

Ivy Tasi @ivymyt

Vitaly Vidmirov @vvid

@ka92
Twitter-driven research: #pix2pix

Brannon Dorsey @brannondorsey

Mario Klingemann @quasimondo

Bertrand Gondouin @bgondouin
“Do as I Do”

OpenPose

pix2pix
Everybody Dance Now

Caroline Chan, Shiry Ginosar, Tinghui Zhou, Alexei A. Efros
UC Berkeley
Results

https://www.youtube.com/watch?v=PCBTZh41Ris&feature=youtu.be
Paired training examples

\{ \text{handbag}, \text{purse} \}

\{ \text{backpack}, \text{purse} \}

\{ \text{backpack}, \text{purse} \}

\ldots

Unpaired training examples

\[ X \]

\{ \text{horse}, \text{horse}, \text{horse} \}

\{ \text{horse}, \text{horse}, \text{horse} \}

\ldots

\[ Y \]

\{ \text{zebras}, \text{zebras}, \text{zebras} \}

\{ \text{zebras}, \text{zebras}, \text{zebras} \}

\ldots
CycleGAN, or “there and back aGAN” [Zhu*, Park*, Isola, Efros. ICCV 2017]
Cycle-Consistency Loss

\[ \|F(G(x)) - x\|_1 \]
Cycle-Consistency Loss

$\mathbf{x}$ \rightarrow G(\mathbf{x}) \rightarrow F(G(\mathbf{x})) \rightarrow y \rightarrow F(y) \rightarrow G(F(y)) \rightarrow \hat{\mathbf{y}}$

- $D_Y$ and $D_X$ are discriminators.
- $\|F(G(\mathbf{x})) - \mathbf{x}\|_1$ and $\|G(F(y)) - y\|_1$ are reconstruction errors.

Reconstruction error

Video
Collection Style Transfer

Photograph © Alexei Efros

Monet

Van Gogh

Cezanne

Ukiyo-e
CG to Real

Grand Theft Auto
Shallower depth of field
Failure case