Image restoration with GANs

Contents

- various image degradations
- recap of GANs
- objective functions
- metrics, perception-distortion trade-off
- DeblurGAN-v2 - ICCV’19 accepted paper on deblurring via GANs
- problems with real-world data setting
- other GAN-based works on image restoration at latest conferences

Image Degradations

Image Degradations

- **blur**

Image Degradations

- blur
- **haze**

Image Degradations

- blur
- haze
- **noise**

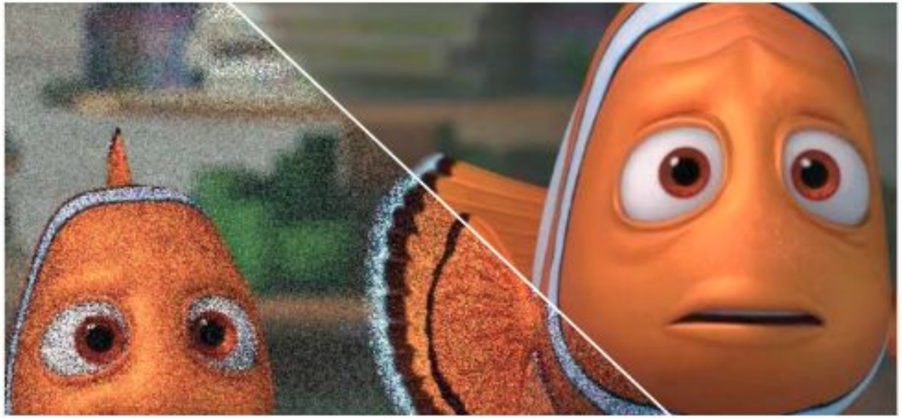


Image Degradations

- blur
- haze
- noise
- **raindrops**

Image Degradations

- blur
- haze
- noise
- raindrops
- **bad lighting**

Image Degradations

- blur
- haze
- noise
- raindrops
- bad lighting
- **low resolution**

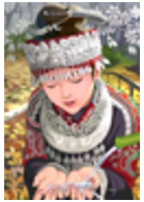


Image Degradations

- blur
- haze
- noise
- raindrops
- bad lighting
- low resolution

and many others

Why do we need this?

- autonomous driving
- video surveillance and security systems
- enhancing media content
- pre-processing to prepare data for other CV tasks

and others

General formulation

$$I_{out} = f\left(I_{gt}\right) + g$$

degradation mapping

additional noise

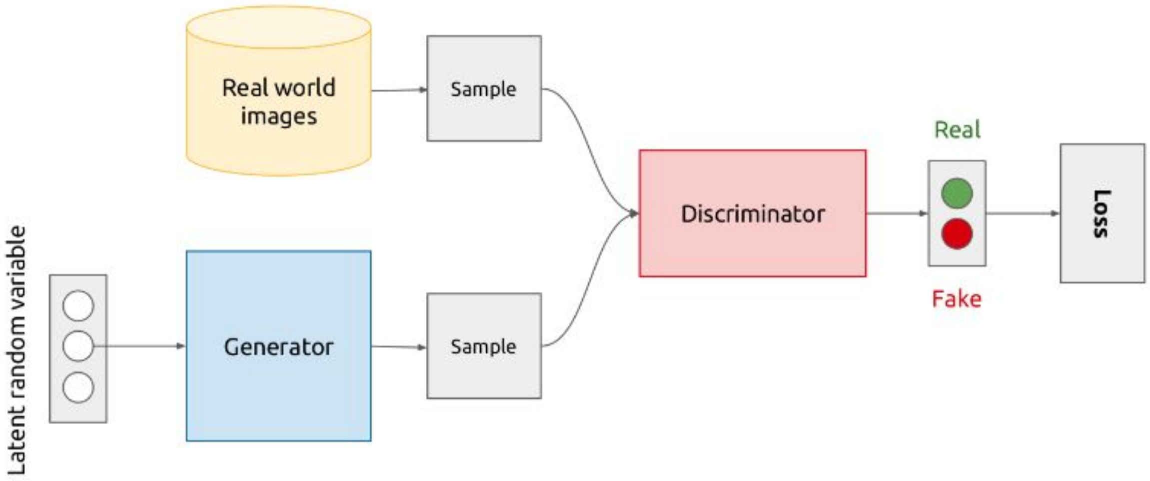
If we don't know the nature of the degradation function, finding the inverse is an ill-posed problem.

Approaches

- non-blind - the mapping is known
- blind - with additional constraints
- completely blind (kernel-free)

How to restore the lost information?

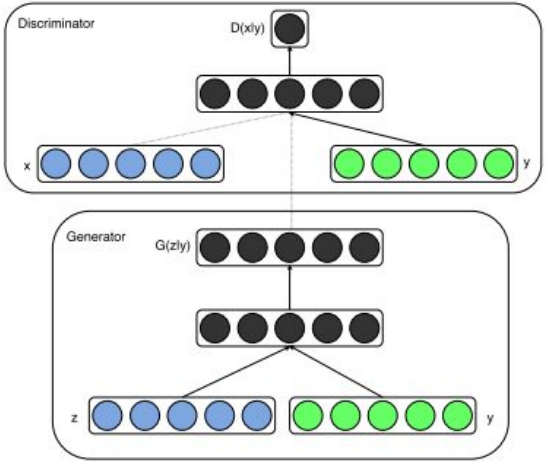
GANs



$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))]$$

https://www.slideshare.net/xavigiro/deep-learning-for-computer-vision-generative-models-and-adversarial-training-upc-2016

Conditional GANs



$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x}|\boldsymbol{y})] + \mathbb{E}_{\boldsymbol{z}\sim p_z(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z}|\boldsymbol{y})))]$$

Conditional Generative Adversarial Nets, Mirza et al, 2014

pix2pix

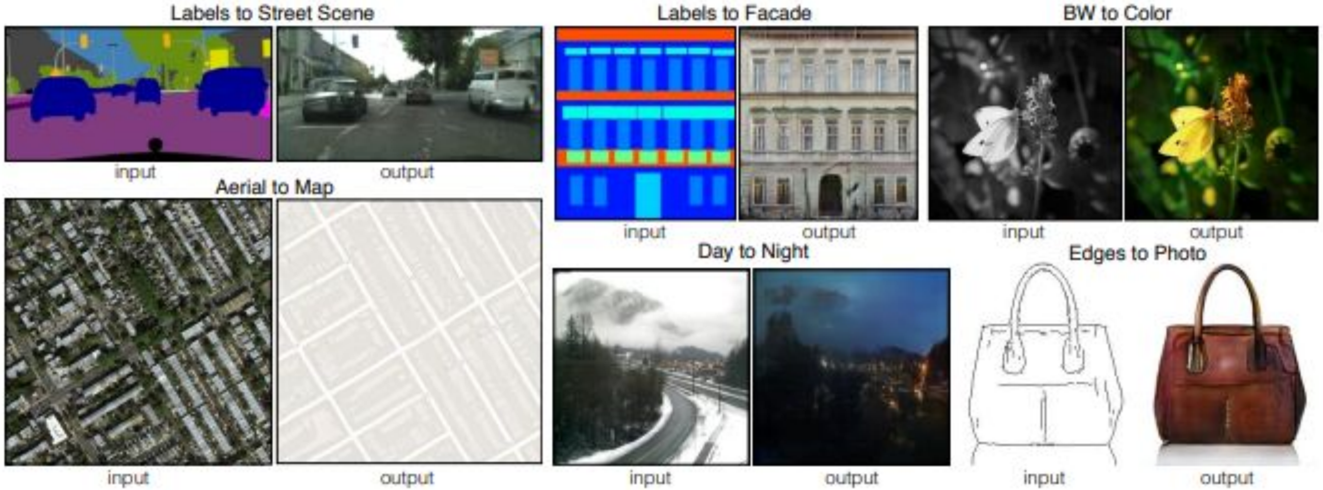


Image-to-Image Translation with Conditional Adversarial Networks, Isola et al, CVPR 2017

WGAN

Kullback-Leibler divergence

$$KL(\mathbb{P}_r||\mathbb{P}_\theta) = \int \log\left(\frac{P_r(x)}{P_\theta(x)}\right) P_r(x) d\mu$$

Earth-Mover distance

$$W(\mathbb{P}_r, \mathbb{P}_\theta) = \inf_{\gamma \in \Pi(\mathbb{P}_r, \mathbb{P}_\theta)} \mathbb{E}_{(x,y)\sim\gamma}[||x - y||]$$

(using Kantorovich-Rubinstein duality)

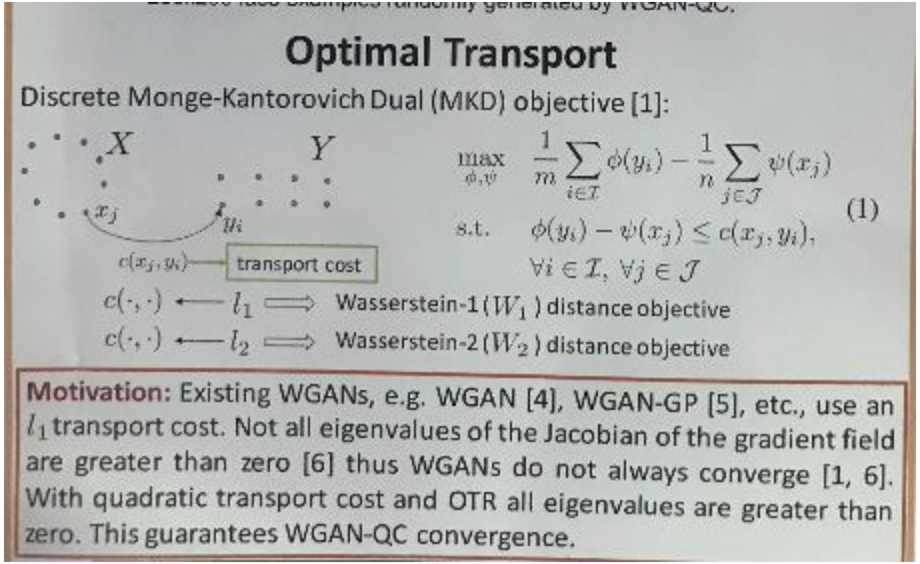
$$W(\mathbb{P}_r, \mathbb{P}_\theta) = \sup_{||f||_L \le 1} \mathbb{E}_{x\sim\mathbb{P}_r}[f(x)] - \mathbb{E}_{x\sim\mathbb{P}_\theta}[f(x)]$$

WGAN-GP

$$L = \mathbb{E}_{x\sim\mathbb{P}_r}[f(x)] - \mathbb{E}_{x\sim\mathbb{P}_\theta}[f(x)] + \lambda \mathbb{E}_{\hat{x}\sim\mathbb{P}_{\hat{x}}}\left[(||\nabla_{\hat{x}} f(\hat{x})||_2 - 1)^2\right]$$

WGAN-QC

sorry for the poor quality, it's from the poster session :)

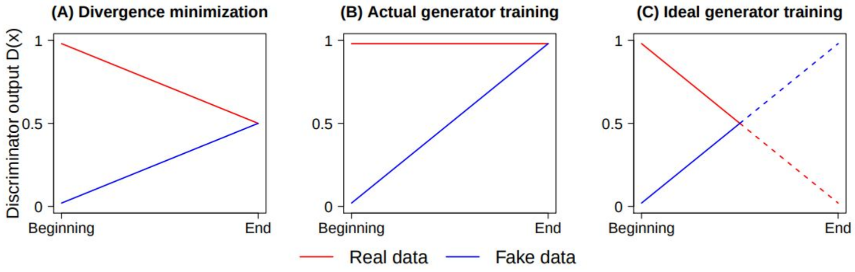


Theorem 2. *Suppose Assumptions 1 and 2 are satisfied, then for small enough learning rate α, there exists λ such that WGAN-QC converges to a local equilibrium point.* [6]

Wasserstein GAN with Quadratic Transport Cost, Liu et al, ICCV 2019

Relativistic approach

“ [...] the probability of real data being real (D(x_r)) should decrease as the probability of fake data being real (D(x_f)) increase”.



$$L_D^{RSGAN} = -\mathbb{E}_{(x_r,x_f)\sim(\mathbb{P},\mathbb{Q})}\left[\log(\text{sigmoid}(C(x_r) - C(x_f)))\right]$$

$$L_G^{RSGAN} = -\mathbb{E}_{(x_r,x_f)\sim(\mathbb{P},\mathbb{Q})}\left[\log(\text{sigmoid}(C(x_f) - C(x_r)))\right]$$

<u>The intuition:</u>
the discriminator estimates the probability that the given real data is more realistic than a randomly sampled fake data.

The relativistic discriminator: a key element missing from standard GAN, Jolicoeur-Martineau

Traditional pixel-based loss issues

- responsible for low-level consistency
- responsible for preserving colours

BUT

- favours for blurry outputs

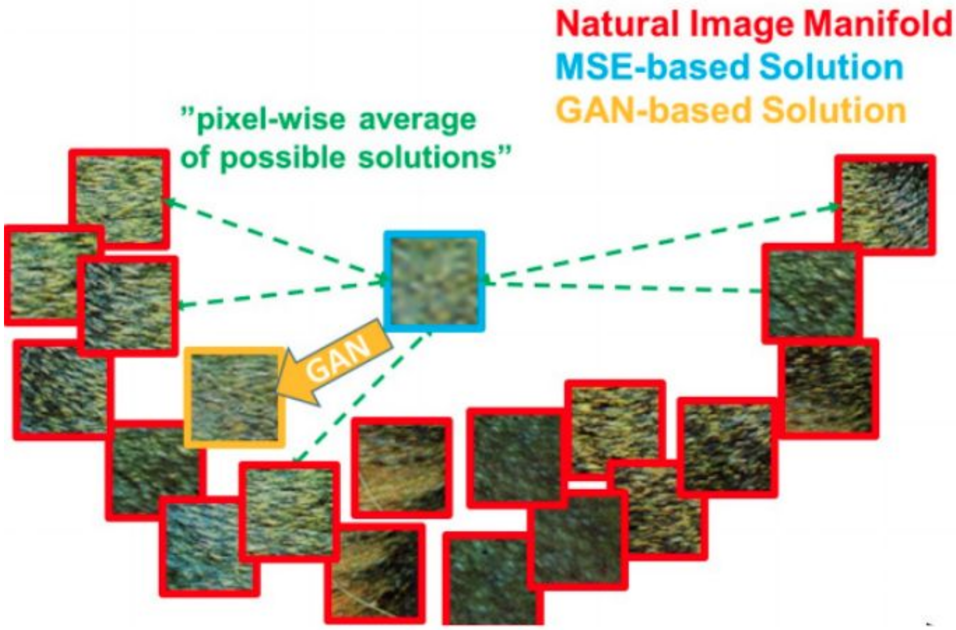
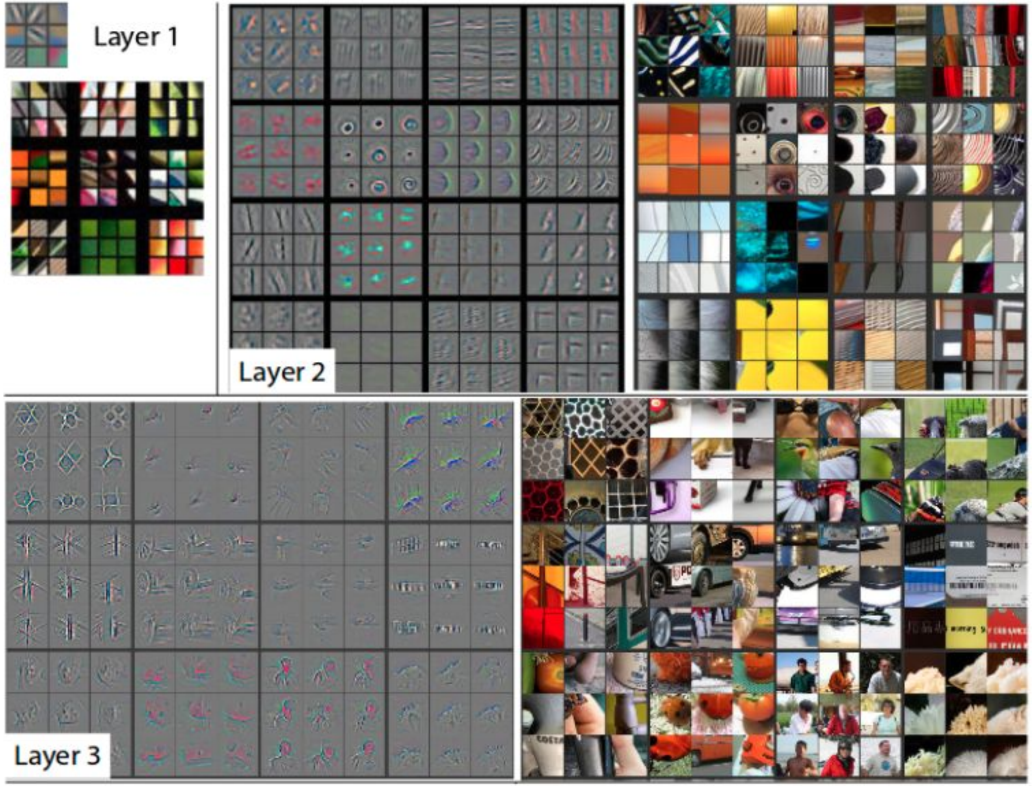


Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, Ledig et al.

Perceptual loss

- pretrained VGG-19
- MSE of features, not raw pixels
- restores general content

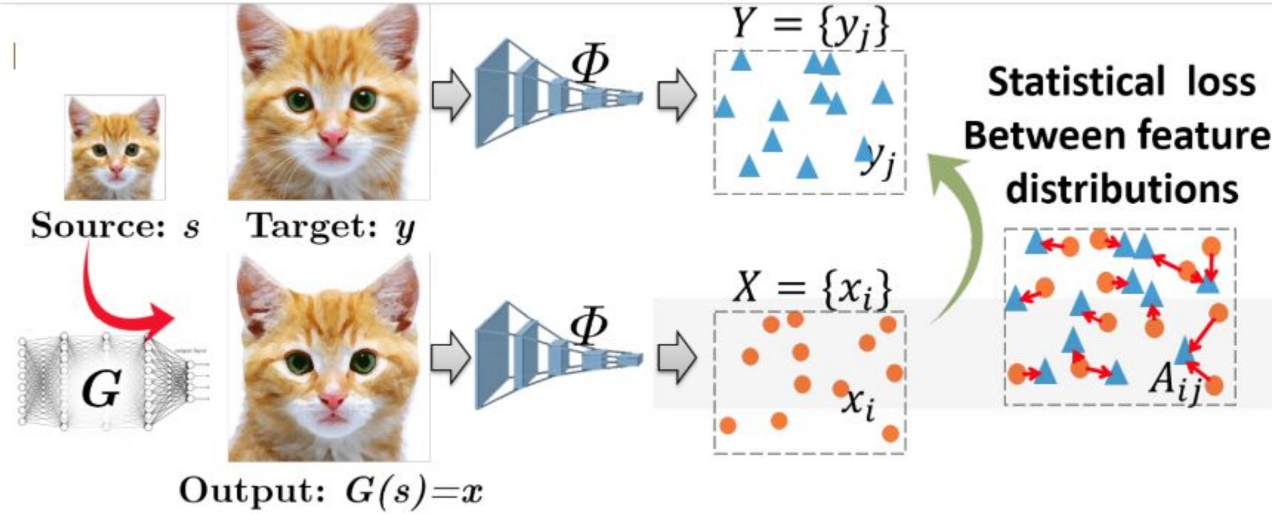


Issues with image semantics understanding

style

content

Contextual loss



$$\mathcal{L}_{\mathrm{CX}}(x, y, l) = -\log\left(\mathrm{CX}\left(\Phi^l(x), \Phi^l(y)\right)\right)$$

The Contextual Loss for Image Transformation with Non-Aligned Data, Mechrez et al.

Traditional problems

- mode collapse
- non-convergence

What can one do?

- **feature matching** - we train the generator to match the expected value of the features on an intermediate layer of the discriminator.
- **minibatch discrimination** - discriminator can compare samples across the batch
- **one-sided label smoothing** - smoothing only positive labels
- **virtual batch normalization** - fixed reference batch
- **adding noise** to the training data
- **balancing updates** of the generator and discriminator
- playing around with different **loss components and hyperparameters**
- **“unrolling” updates** of discriminator
- using **multiple GANs**

and many more

Improved Techniques for Training GANs, Salimans et al.

How do we evaluate the results?

Metrics, perception-distortion trade-off

- PSNR

$$PSNR = 20 \log_{10} \frac{MAX_g}{\sqrt{MSE}}$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [g(i,j) - \tilde{g}(i,j)]^2$$

Metrics, perception-distortion trade-off

- PSNR

$$PSNR = 20 \log_{10} \frac{MAX_g}{\sqrt{MSE}}$$

- SSIM

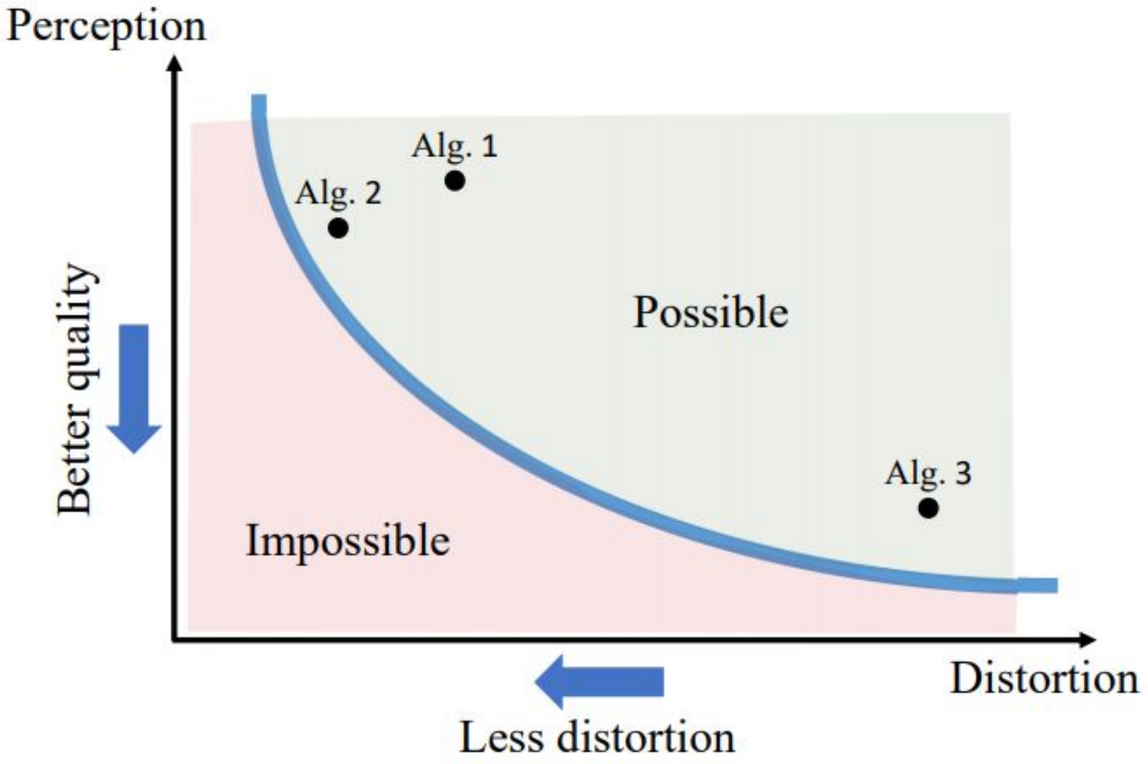
$$SSIM(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^{\alpha} [c(\mathbf{x}, \mathbf{y})]^{\beta} [s(\mathbf{x}, \mathbf{y})]^{\gamma}$$

luminance - $$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

contrast - $$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

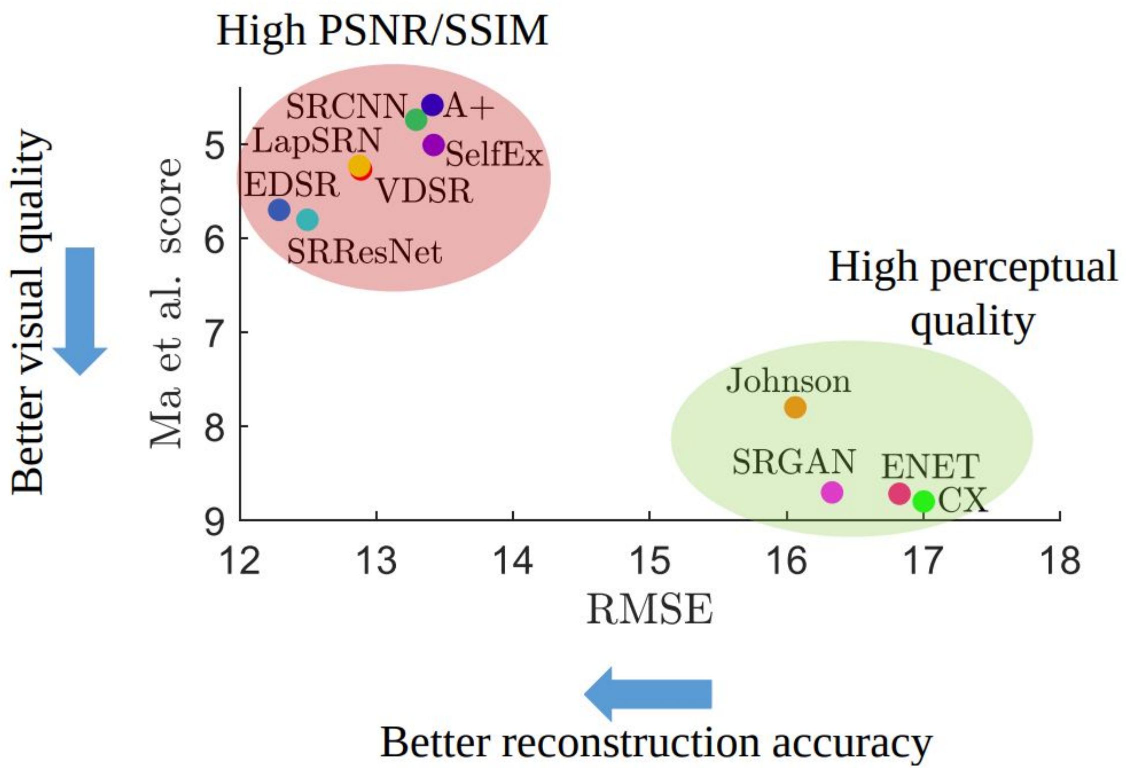
structure - $$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}$$

Metrics, perception-distortion trade-off



The Perception-Distortion Tradeoff, Blau et al., CVPR 2018

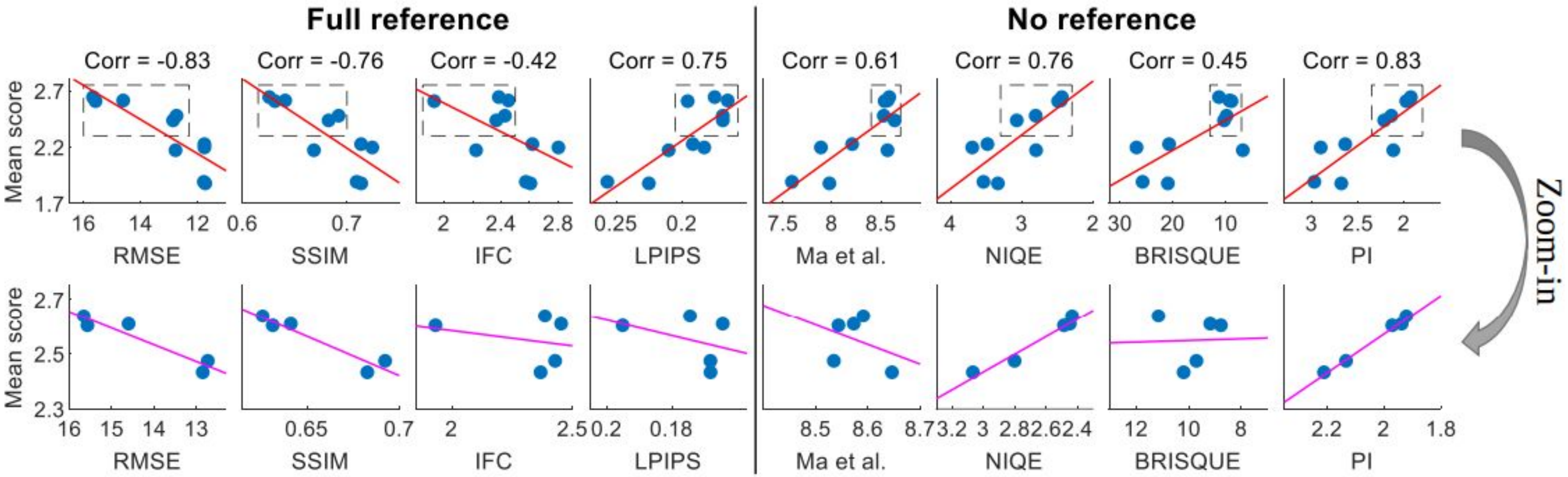
Metrics, perception-distortion trade-off



The 2018 PIRM Challenge on Perceptual Image Super-resolution

Metrics, perception-distortion trade-off

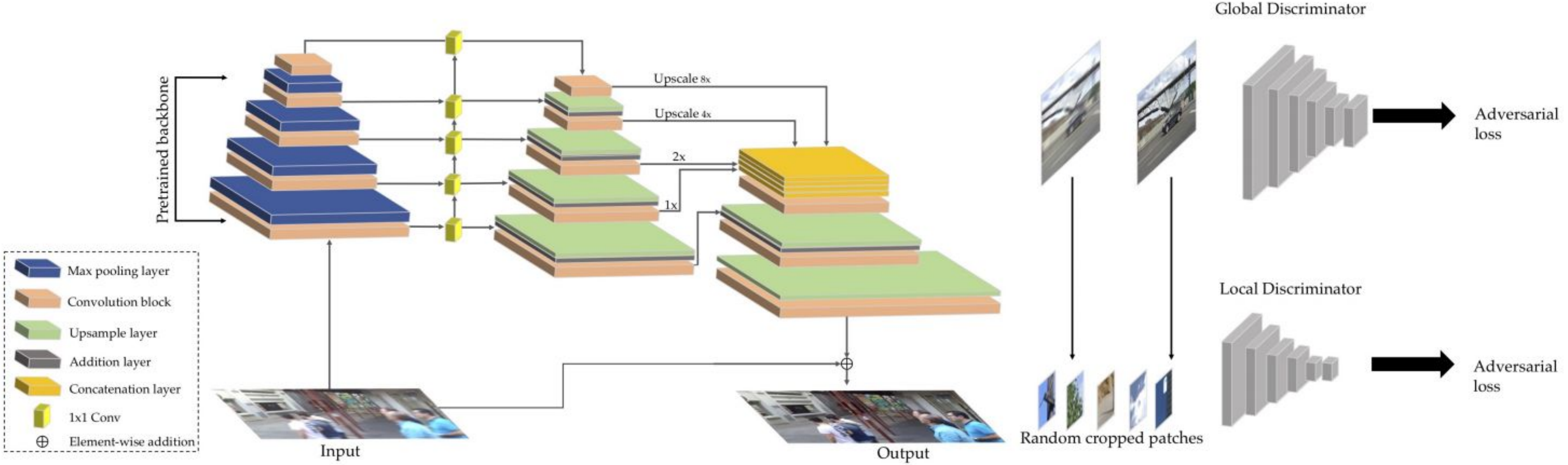
$$\mathrm{PI} = \frac{1}{2}\left((10 - \mathrm{Ma}) + \mathrm{NIQE}\right)$$



The 2018 PIRM Challenge on Perceptual Image Super-resolution, Blau et al, ECCV 2018

Moving to latest results

DeblurGAN-v2



DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better, Kupyn et al., ICCV 2019

DeblurGAN-v2

Blurry	Krishnan *et al.* [20]	Whyte *et al.* [49]	Xu *et al.* [51]	Sun *et al.* [43]	Pan *et al.* [36]
1	1.08	0.57	0.77	0.64	0.91

DeepDeblur [33]	SRN [45]	DeblurGAN [21]	DeblurGAN-v2 (Inception-ResNet-v2)	DeblurGAN-v2 (MobileNet)	DeblurGAN-v2 (MobileNet-DSC)
1.08	1.68	1.29	**1.74**	1.44	1.32

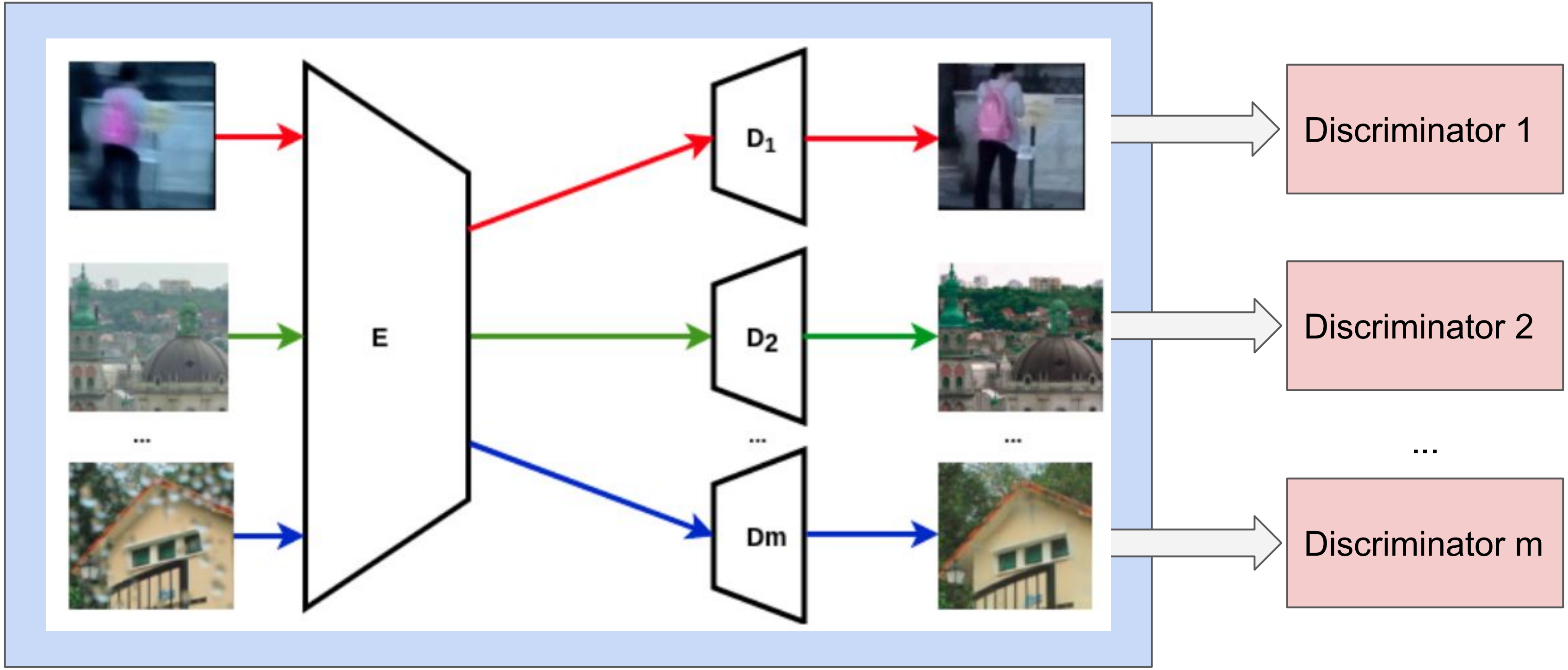
(a) Blurred photo (b) Whyte *et al.* [49] (c) Krishnan *et al.* [20] (d) Sun *et al.* [43]

(e) Xu *et al.* [51] (f) Pan *et al.* [36] (g) DeepDeblur [33] (h) SRN [45]

(i) DeblurGAN [21] (j) DeblurGAN-v2 (Inception-ResNet-v2) **[Best visual quality]** (k) DeblurGAN-v2 (MobileNet) **[High efficiency]** (l) DeblurGAN-v2 (MobileNet-DSC) **[Highest efficiency]**

DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better, Kupyn et al., ICCV 2019

Multi-task learning pipeline



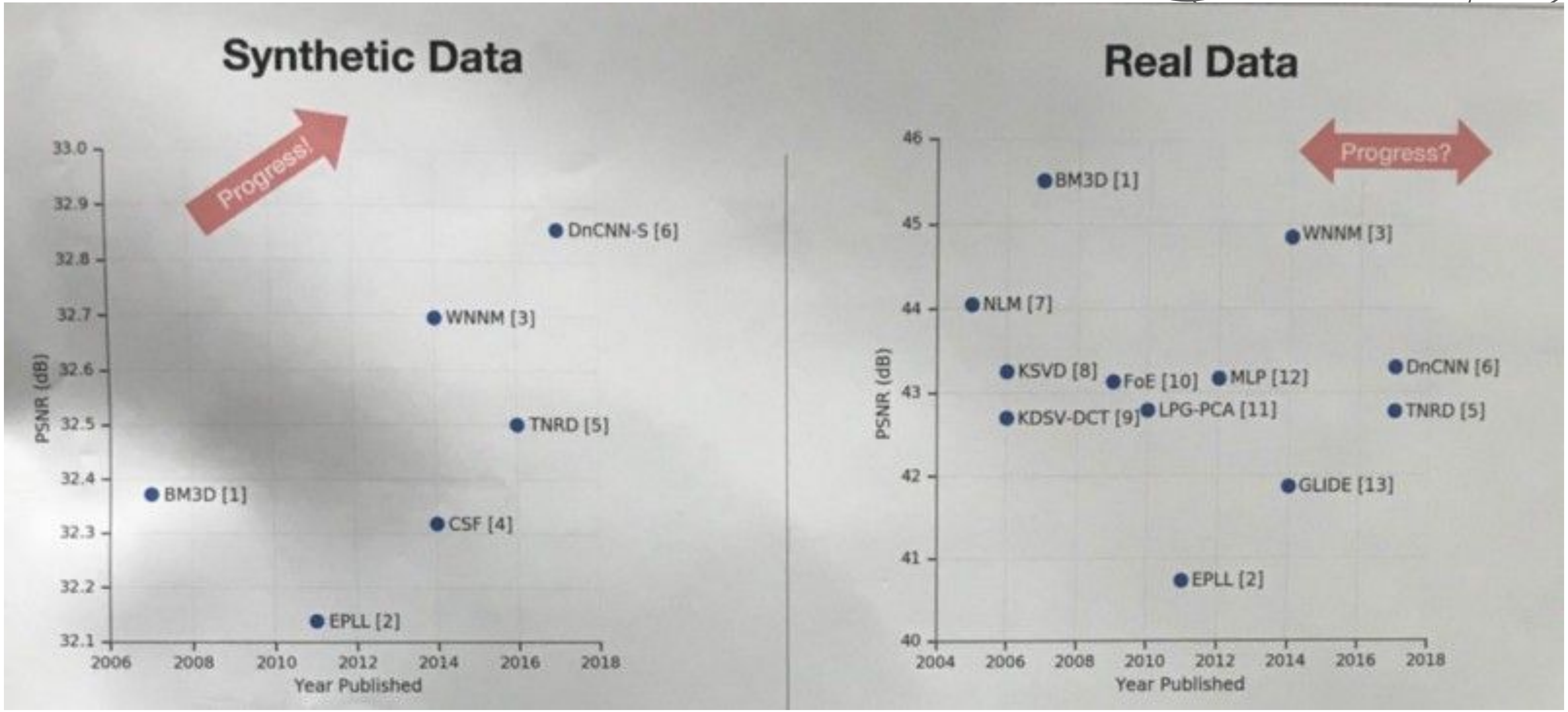
Multi-task Generator

Synthetic data

DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better, Kupyn et al., ICCV 2019

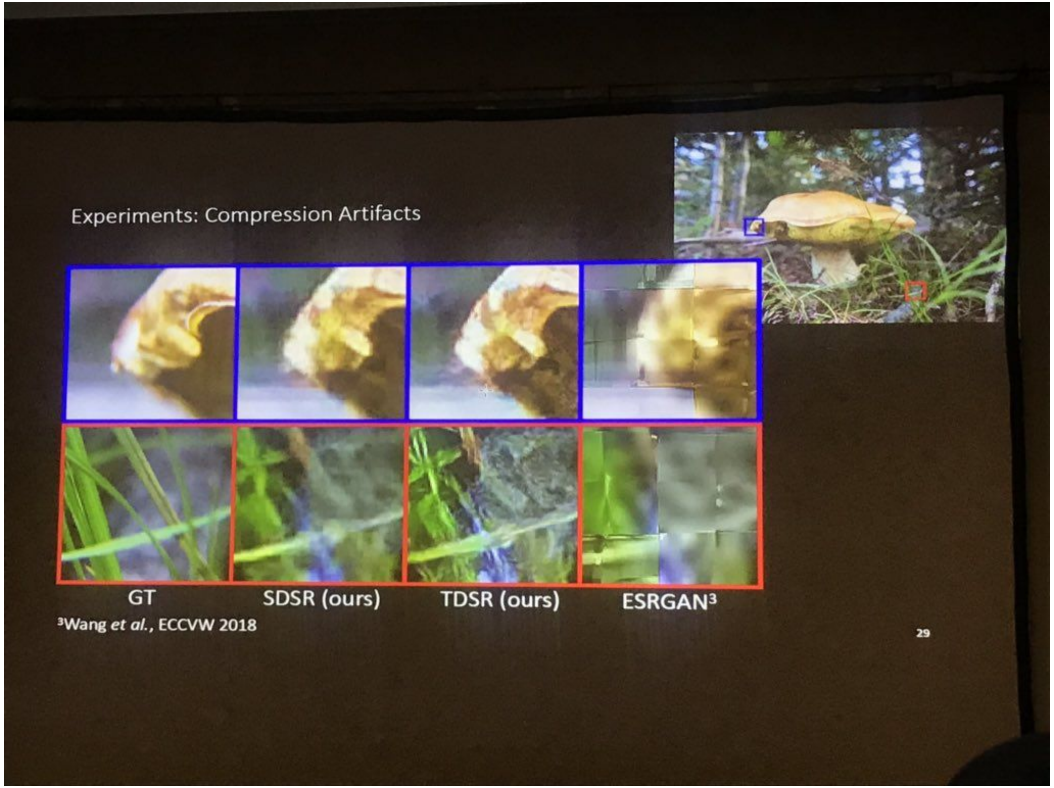
Synthetic data - denoising

sorry for the poor quality,
it's from the poster session :)



Unprocessing Images for Learned Raw Denoising, Brooks et al., CVPR 2019

Synthetic data - super-resolution

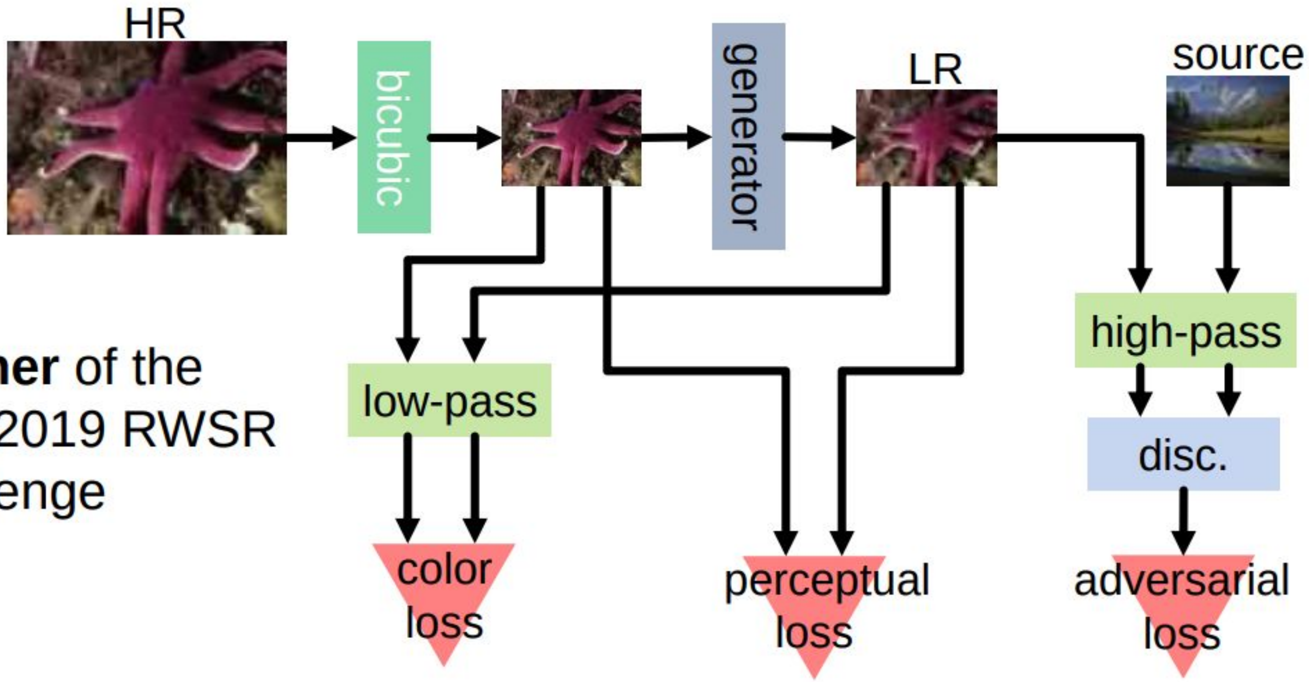


Frequency separation for Real-World SR, Fritsche et al., ICCVW 2019

Synthetic data - super-resolution

https://github.com/xinntao/ESRGAN

Frequency Separation for Real-World SR



Winner of the AIM 2019 RWSR challenge

Frequency separation for Real-World SR, Fritsche et al., ICCVW 2019

Lack of real data

Model-based Data Generation

Aim to accurately model the image generation process

- Blur kernel, noise, compression, etc.

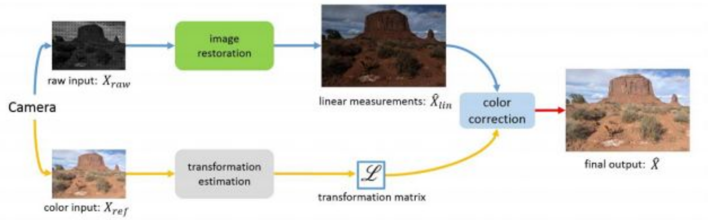
☺ More realistic data

☺ Use sensor specific info

☹ Often hard

☹ Sensor/dataset specific

☹ Often limited knowledge
of the image formation process



Towards Real Scene Super-Resolution with Raw Images. Xiangyu Xu, Yongrui Ma, Wenxiu Sun. CVPR 2019.

http://www.vision.ee.ethz.ch/fire19/speakers/FIRE_tutorial_talk2_MartinDanelljan.pdf

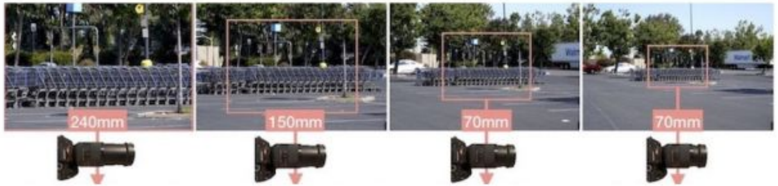
Lack of real data

Collect Real Data

Methods for collecting real data

☺ Real data
☺ No models required

☹ Specific setups
☹ Cumbersome and expensive
☹ Limitations (misalignments, distance, lightning)

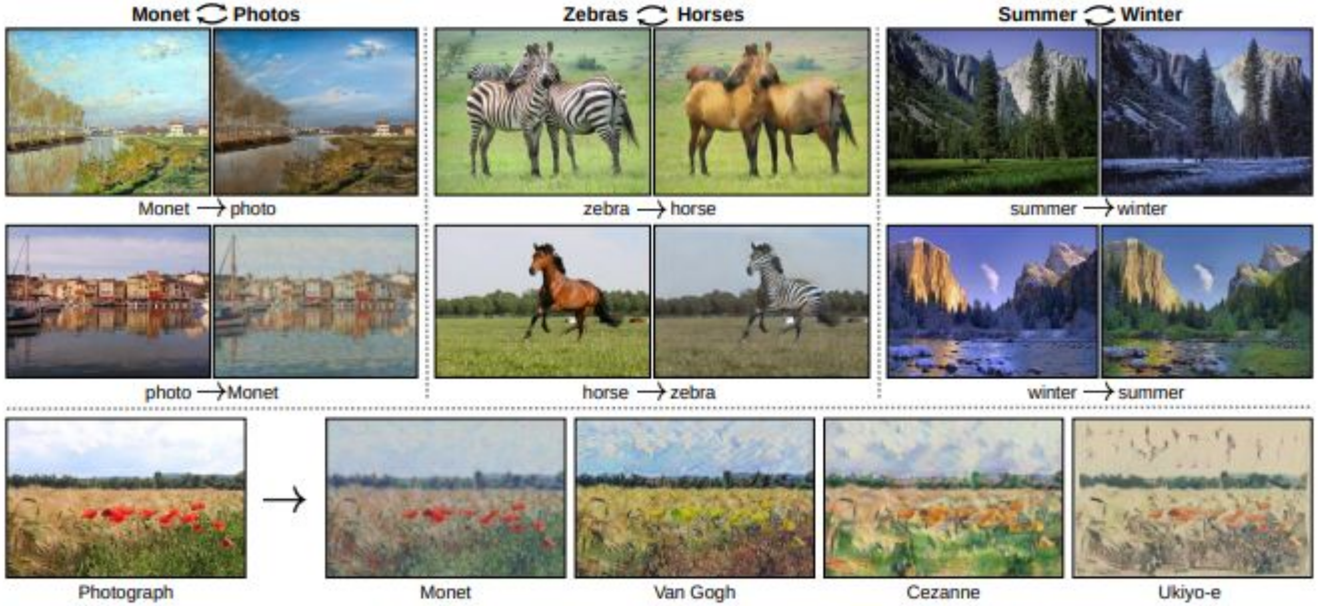


Zoom to Learn, Learn to Zoom. Xuaner Zhang, Qifeng Chen, Ren Ng, Vladlen Koltun. CVPR 2019.

Camera Lens Super-Resolution. Chang Chen, Zhiwei Xiong, Xinmei Tian, Zheng-Jun Zha, Feng Wu. CVPR 2019.

http://www.vision.ee.ethz.ch/fire19/speakers/FIRE_tutorial_talk2_MartinDanelljan.pdf

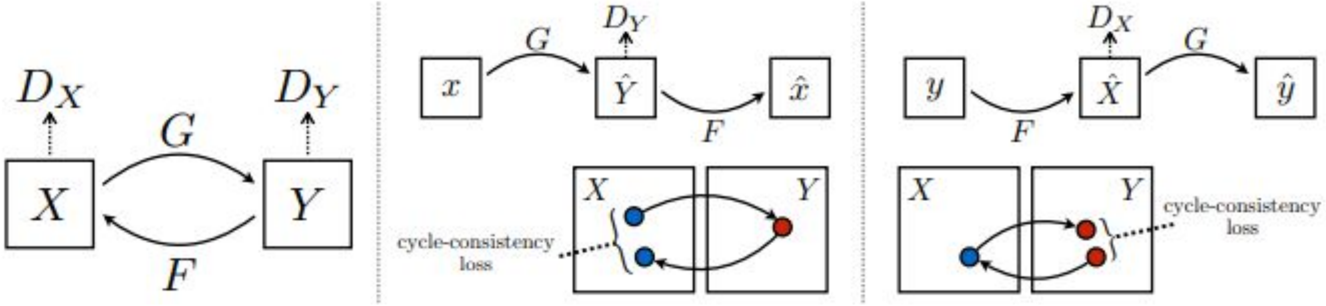
CycleGAN



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

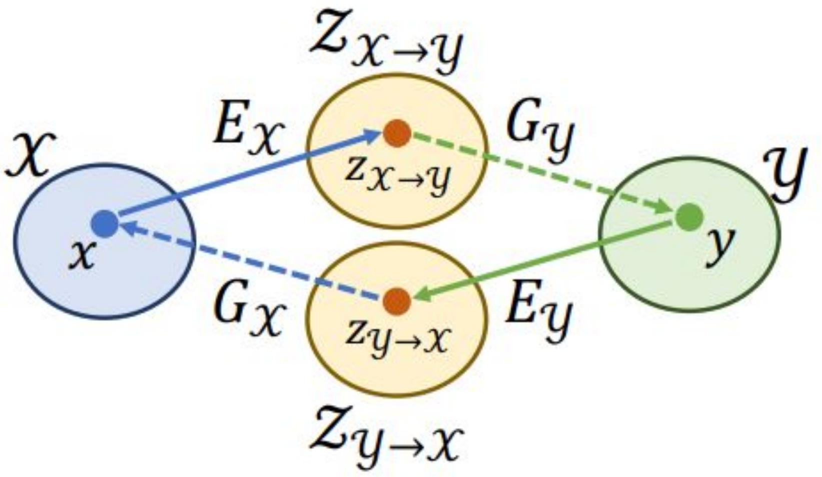
CycleGAN

Cycle-consistency

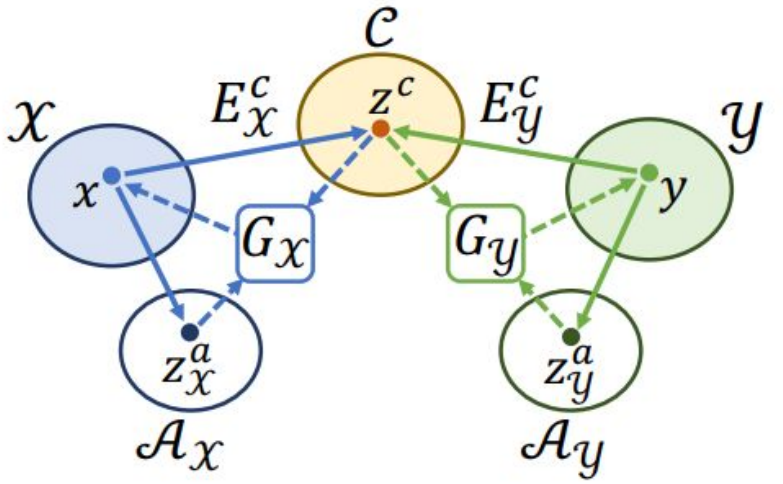


Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, Zhu et al, ICCV 2017

Improvements to CycleGAN: DRIT



CycleGAN

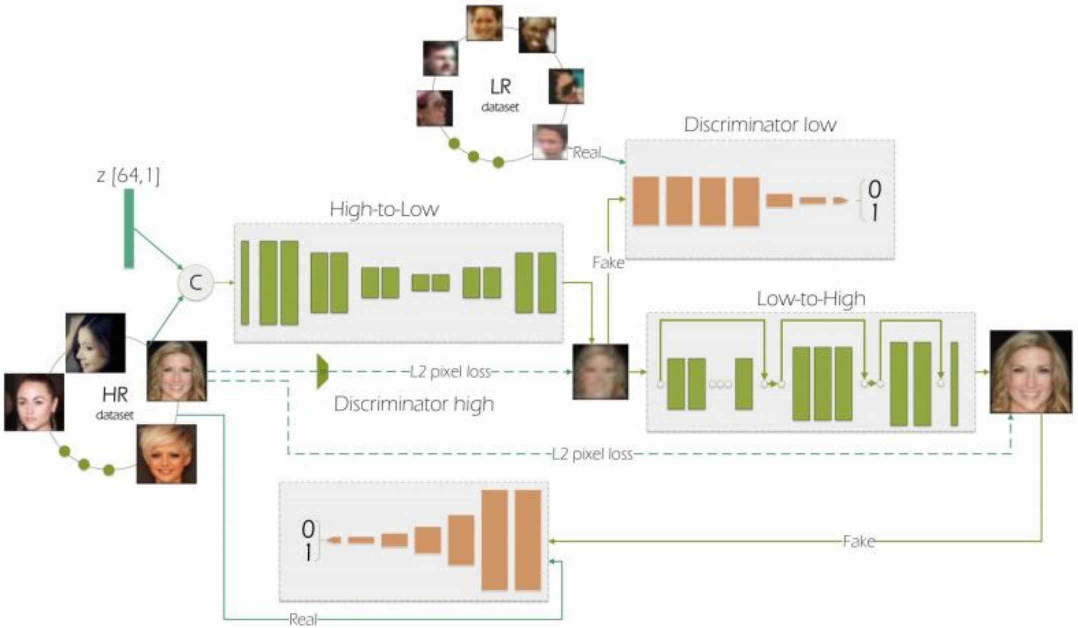


DRIT++

Diverse Image-to-Image Translation via Disentangled Representations, Lee et al, ECCV 2018

Use a GAN to Learn Image Degradation

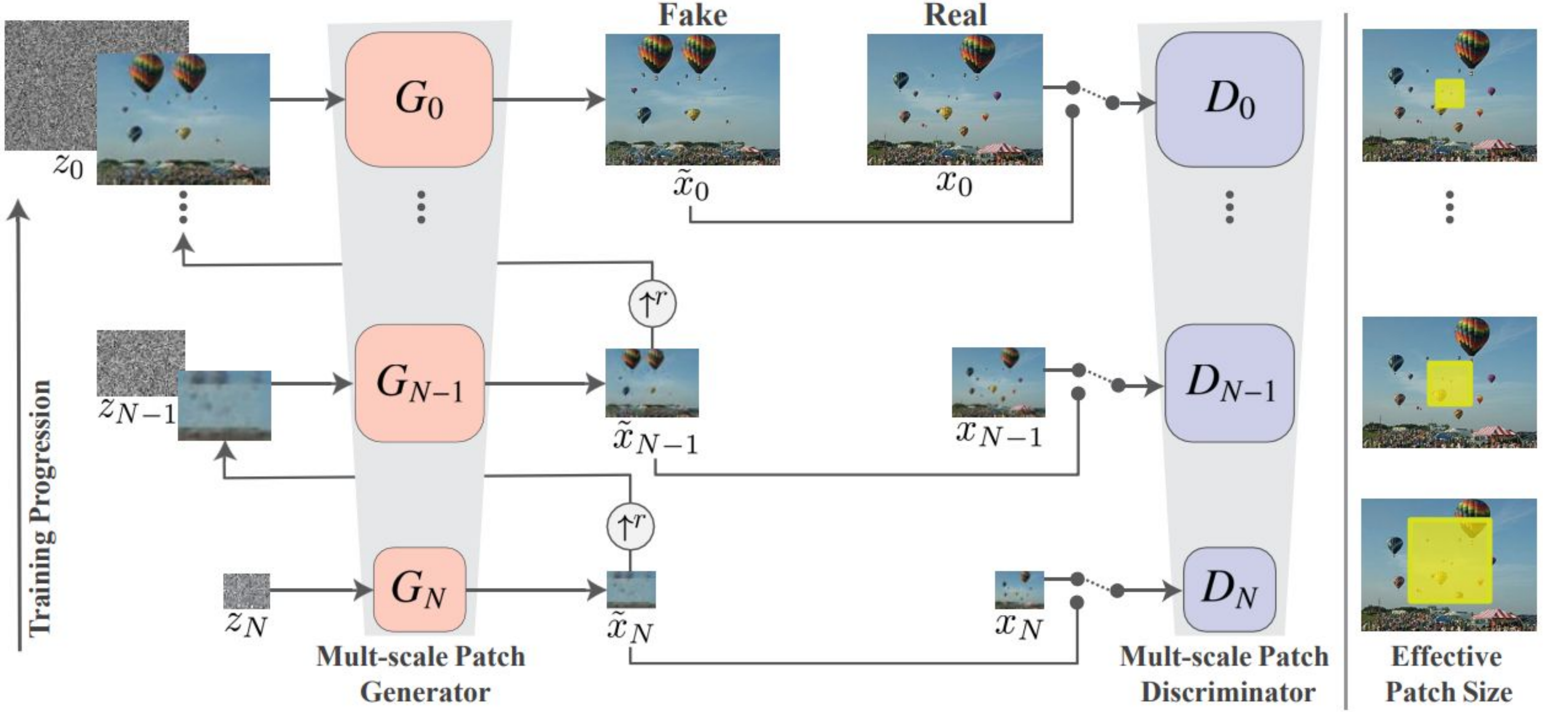
- Faces
- Learn HR->LR net
- Joint fine-tuning of full cycle



To learn image super-resolution, use a GAN to learn how to do image degradation first, Bulat et al, ECCV 2018

Fresh from the oven

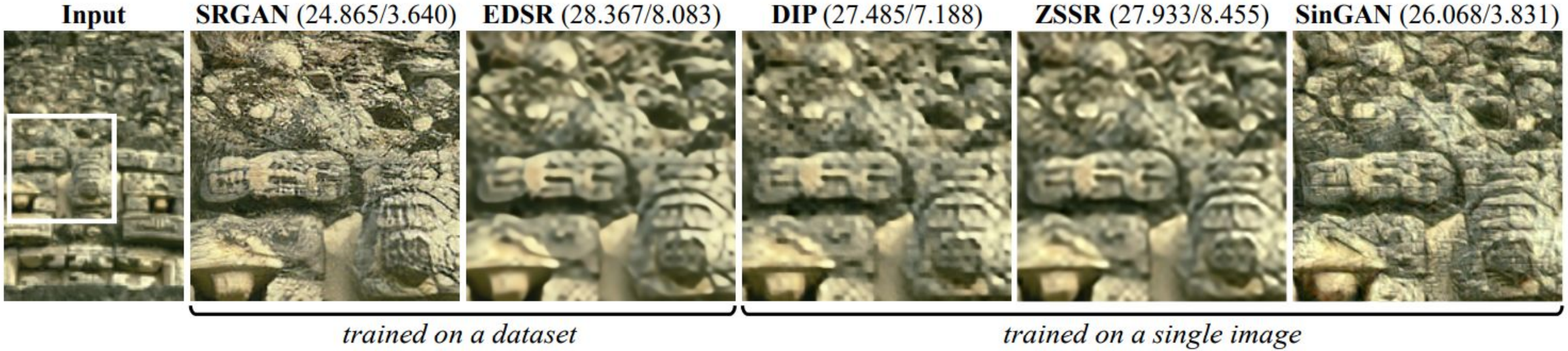
SinGAN



SinGAN: Learning a Generative Model from a Single Natural Image, TR Shaham et al., ICCV 2019 Best Paper

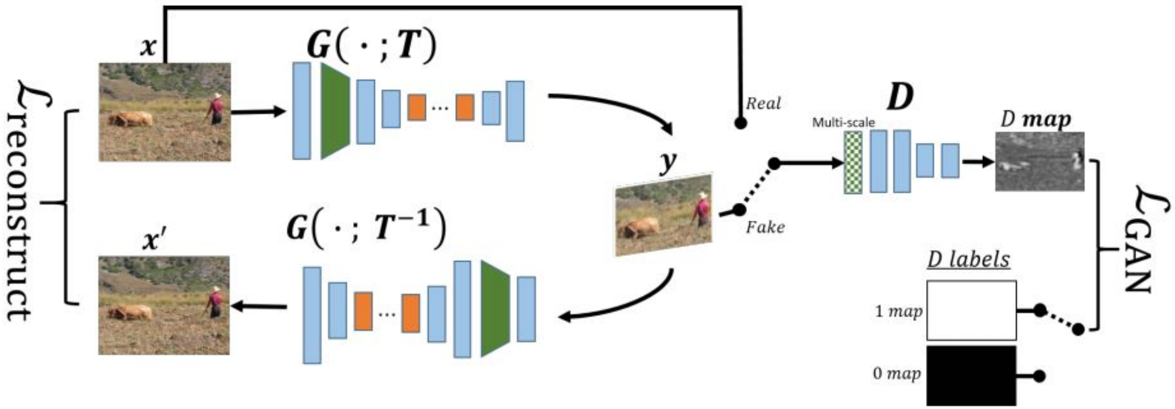
SinGAN

Super-resolution



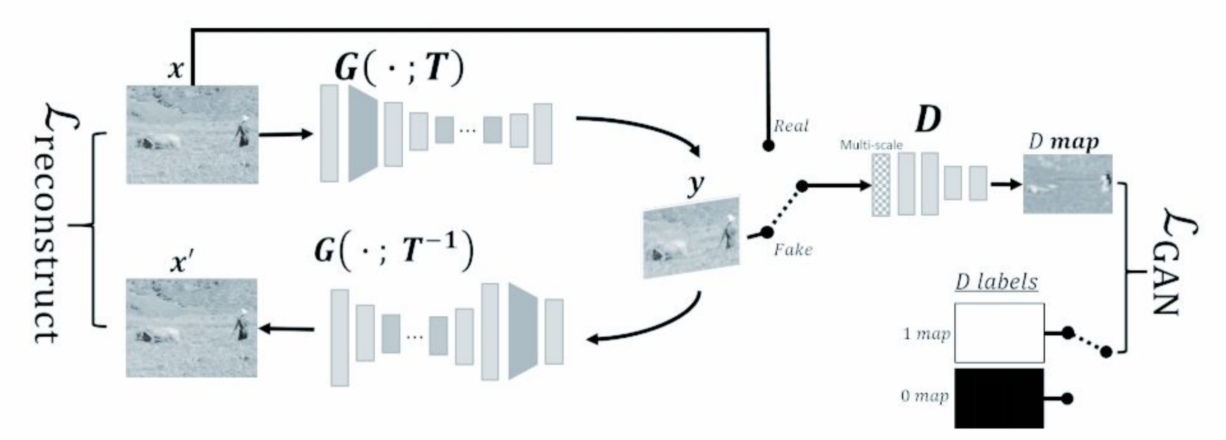
SinGAN: Learning a Generative Model from a Single Natural Image, TR Shaham et al., ICCV 2019 Best Paper

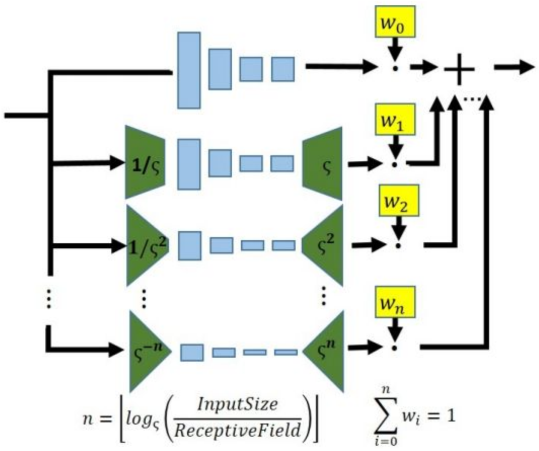
InGAN



InGAN: Capturing and Remapping the "DNA" of a Natural Image, Shocher et al., ICCV 2019

InGAN





Adaptive Multi-scale
Patch Discriminator

InGAN: Capturing and Remapping the "DNA" of a Natural Image, Shocher et al., ICCV 2019

InGAN

InGAN: Capturing and Remapping the “DNA” of a Natural Image, Shocher et al., ICCV 2019

Thank you! Questions?