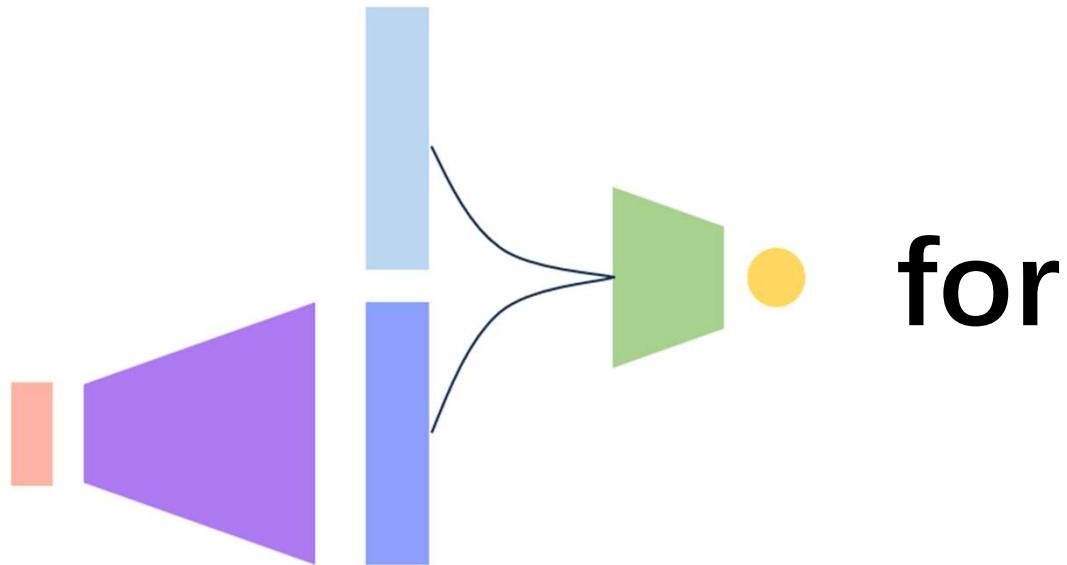


基于生成网络的数字图像处理

GAN for Digital Image Processing

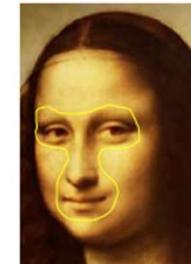


Various DIP tasks



Geometric Editing

Various DIP tasks



source/destination



cloning



seamless cloning

Color Space Editing

Various DIP tasks



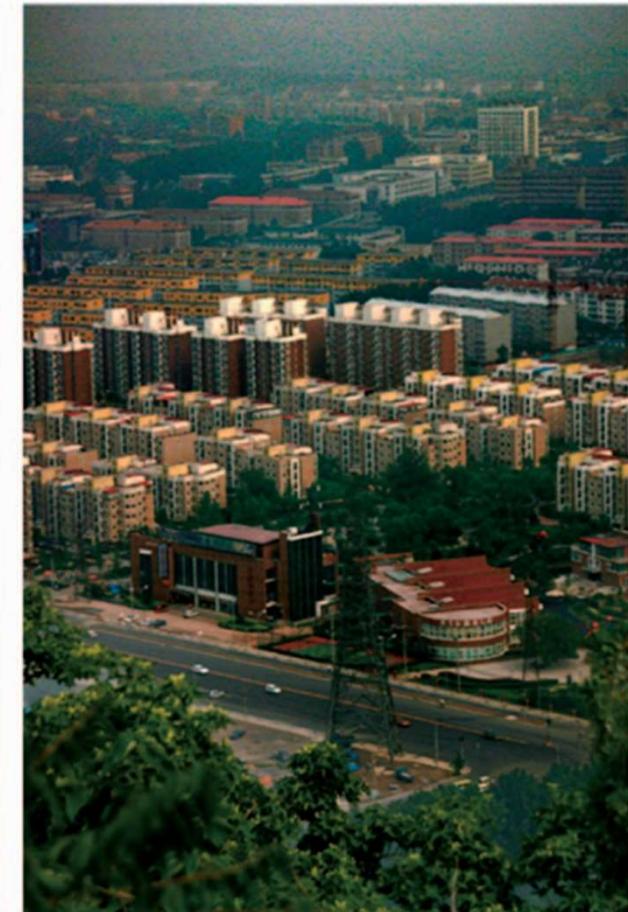
Image Segmentation

Various DIP tasks



Image Enhancement

Various DIP tasks



Haze Removal

Various DIP tasks

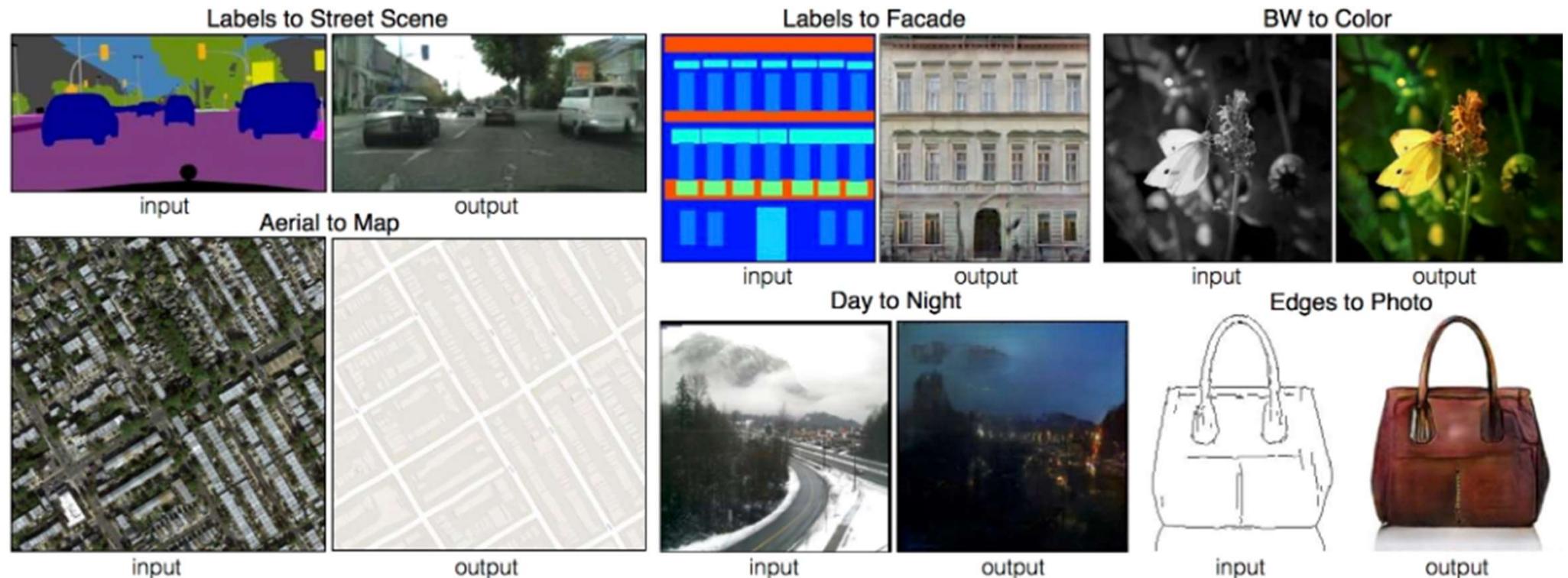


Image to Image

Various DIP tasks



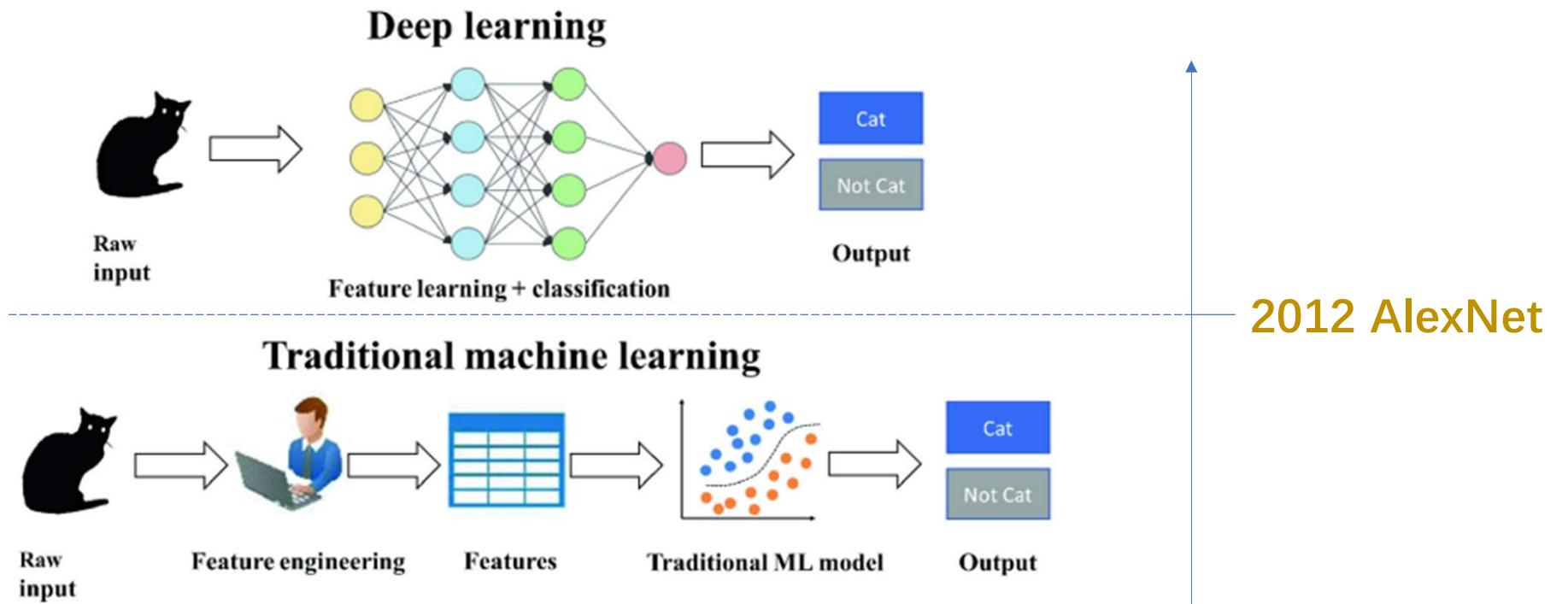
Image Supersolution

Various DIP tasks

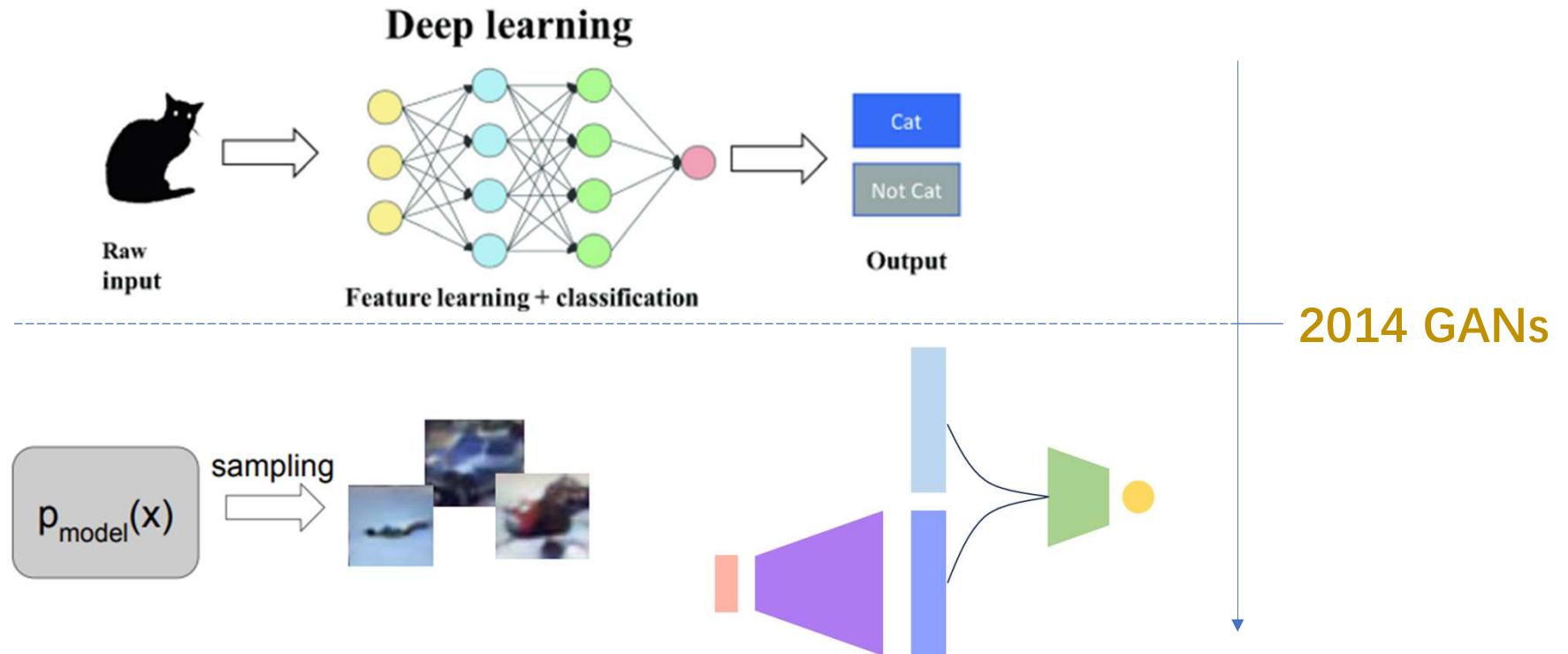


Image Matting

Previous Tools



New Tool



1. GANs能否做这些DIP的任务?
2. 假如能，优势是什么？
3. 为什么可以做这些？（方法论）

GANs能否做这些DIP的任务？— Way 1

我在准备一节课的课件，课程内容是讲生成模型（generative models），尤其是GAN，在数字图像处理（Digital Image Processing）中的应用，你能举出几个GAN用于不同图像处理的任务吗，请越多越好，并给出相关paper



Thought for 46 seconds ▾

生成对抗网络（GAN）在数字图像处理领域有广泛的应用，以下是一些典型任务及相关论文：

1. 图像超分辨率 (Image Super-Resolution)

- 描述：从低分辨率图像生成高分辨率图像，提高图像的细节和清晰度。
- 论文：Christian Ledig 等人，"Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network" (SRGAN) , CVPR 2017。

2. 图像修复 (Image Inpainting)

- 描述：填充和恢复图像中缺失或受损的部分，生成完整的图像。
- 论文：Jiahui Yu 等人，"Generative Image Inpainting with Contextual Attention" , CVPR 2018。

3. 图像去模糊 (Image Deblurring)

- 描述：消除图像中的运动模糊或焦点模糊，提升图像的清晰度。
- 论文：Orest Kupyn 等人，"DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks" , CVPR 2018。

4. 图像着色 (Image Colorization)

- 描述：为灰度图像自动添加颜色，使其转换为彩色图像。
- 论文：Richard Zhang 等人，"Colorful Image Colorization" , ECCV 2016。 (虽然该论文未使用GAN，但GAN也被用于图像着色，如基于Pix2Pix的方法)

5. 图像到图像转换 (Image-to-Image Translation)

- 描述：在不同的图像域之间进行转换，例如将素描转换为照片。
↓
- 论文：Phillip Isola 等人，"Image-to-Image Translation with Conditional Adversarial

7. 图像合成 (Image Synthesis)

- 描述：从随机噪声中生成逼真的图像，如人脸、风景等。
- 论文：Tero Karras 等人，"Progressive Growing of GANs for Improved Quality, Stability, and Variation" , ICLR 2018。

8. 人脸属性编辑 (Facial Attribute Editing)

- 描述：修改人脸图像的特定属性，如年龄、表情、发型等。
- 论文：Yunjey Choi 等人，"StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation" , CVPR 2018。

9. 文本到图像合成 (Text-to-Image Synthesis)

- 描述：根据文本描述生成对应的图像内容。
- 论文：Han Zhang 等人，"StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks" , ICCV 2017。

10. 医学图像合成 (Medical Image Synthesis)

- 描述：在医学领域生成或转换图像，例如从MRI生成CT图像。
- 论文：Hao Dong 等人，"Unsupervised Domain Adaptation for Medical Imaging Segmentation with Self-Ensembling and Adversarial Training" , MICCAI 2018。

GANs能否做这些DIP的任务? — Way 2

The screenshot shows a GitHub repository page for 'gans-awesome-applications'. At the top, there's a navigation bar with 'Watch 208', 'Fork 809', and 'Star 5k'. Below the navigation bar, there are buttons for 'master', 'Branch', 'Tags', 'Go to file', 'Add file', and 'Code'. A search bar is also present. On the left, a sidebar shows the repository structure: 'nashory Merge pull request #17 from hologerry/master' (f034092 · 4 years ago), '33 Commits'. Below this, there are two files listed: 'jpg' (add banner, 7 years ago) and 'README.md' (Add two font generation papers, 4 years ago). To the right, there's an 'About' section with the text 'Curated list of awesome GAN applications and demo' and several tags: 'github', 'applications', 'generative-adversarial-network', 'papers', 'curated-list', and 'demonstration'.

Image Editing

- Invertible Conditional GANs for image editing, [\[paper\]](#), [\[github\]](#)
- Image De-raining Using a Conditional Generative Adversarial Network, [\[paper\]](#), [\[github\]](#)

Face Aging

- Age Progression/Regression by Conditional Adversarial Autoencoder, [\[paper\]](#), [\[github\]](#)
- CAN: Creative Adversarial Networks Generating "Art" by Learning About Styles and Deviating from Style Norms, [\[paper\]](#)
- FACE AGING WITH CONDITIONAL GENERATIVE ADVERSARIAL NETWORKS, [\[paper\]](#)

Human Pose Estimation

- Joint Discriminative and Generative Learning for Person Re-identification, [\[paper\]](#), [\[github\]](#), [\[video\]](#)
- Pose Guided Person Image Generation, [\[paper\]](#)

Domain-transfer (e.g. style-transfer, pix2pix, sketch2image)

Awesome XXX
In github

GANs能否做这些DIP的任务? — Way 3

Explore Topics Trending Collections Events GitHub Sponsors

Generative Adversarial Network

Star

Generative adversarial networks (GAN) are a class of generative machine learning frameworks. A GAN consists of two competing neural networks, often termed the Discriminator network and the Generator network. GANs have been shown to be powerful generative models and are able to successfully generate new data given a large enough training dataset.

Here are 2,995 public repositories matching this topic...



microsoft / Bringing-Old-Photos-Back-to-Life

Star 15.1k

XingangPan / DragGAN

Starred 35.7k

Code Issues Pull requests

Official Code for DragGAN (SIGGRAPH 2023)

artificial-intelligence generative-adversarial-network image-manipulation generative-models

Updated on May 19 Python

svc-develop-team / so-vits-svc

Starred 25.7k

Code Issues Pull requests Discussions

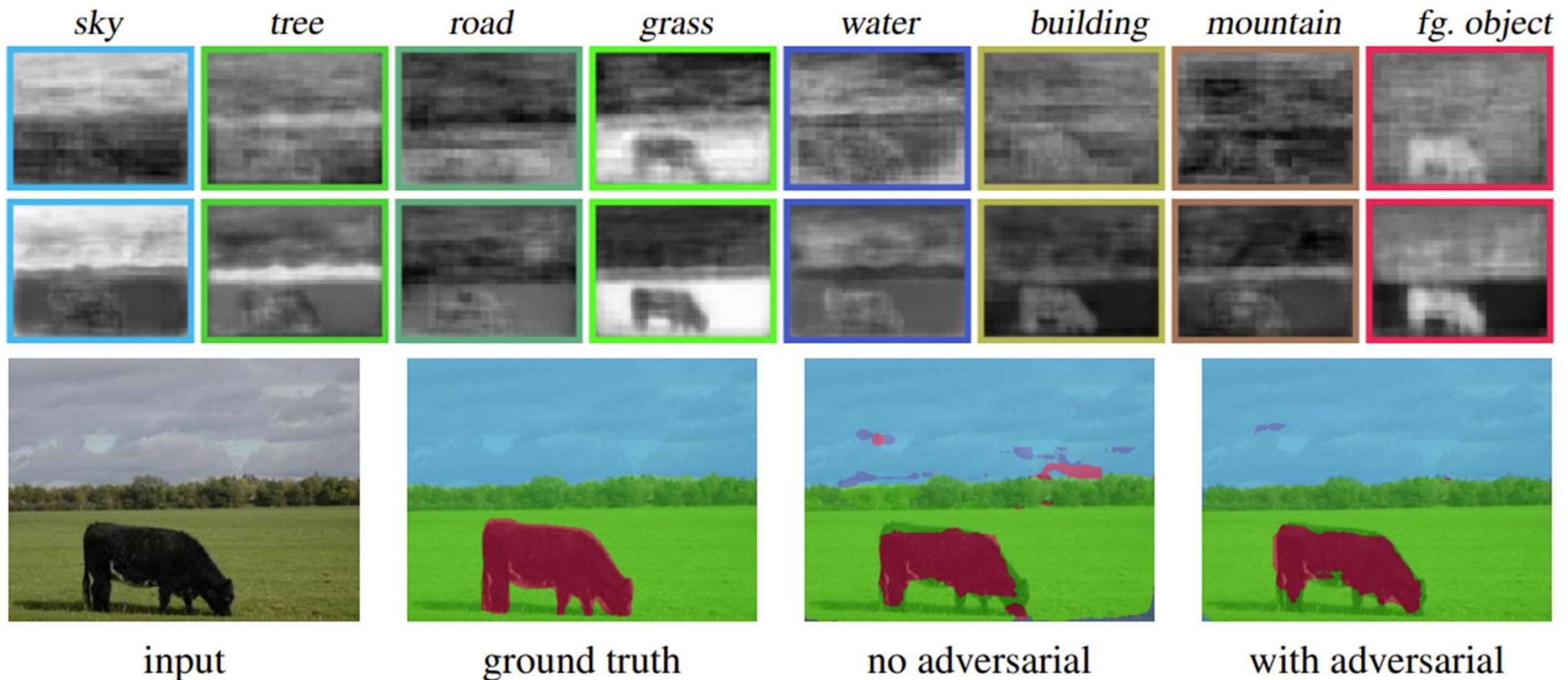
SoftVC VITS Singing Voice Conversion

flow ai deep-learning voice speech pytorch audio-analysis generative-adversarial-network variational-inference voice-conversion vc voice-changer vits singing-voice-conversion voiceconversion sovits so-vits-svc

Updated on Nov 11, 2023 Python

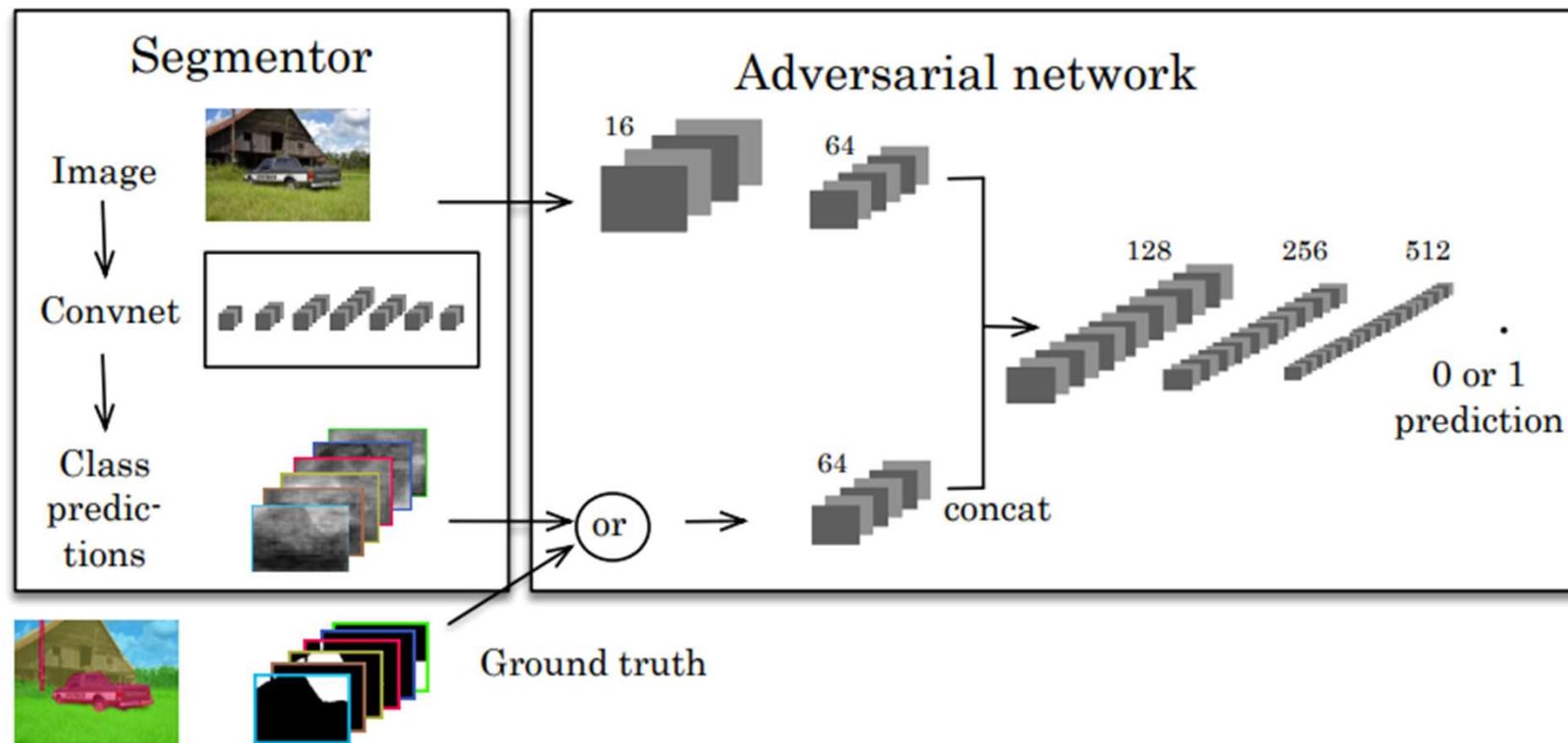
XXX Topics
In github

GANs for Image Segmentation



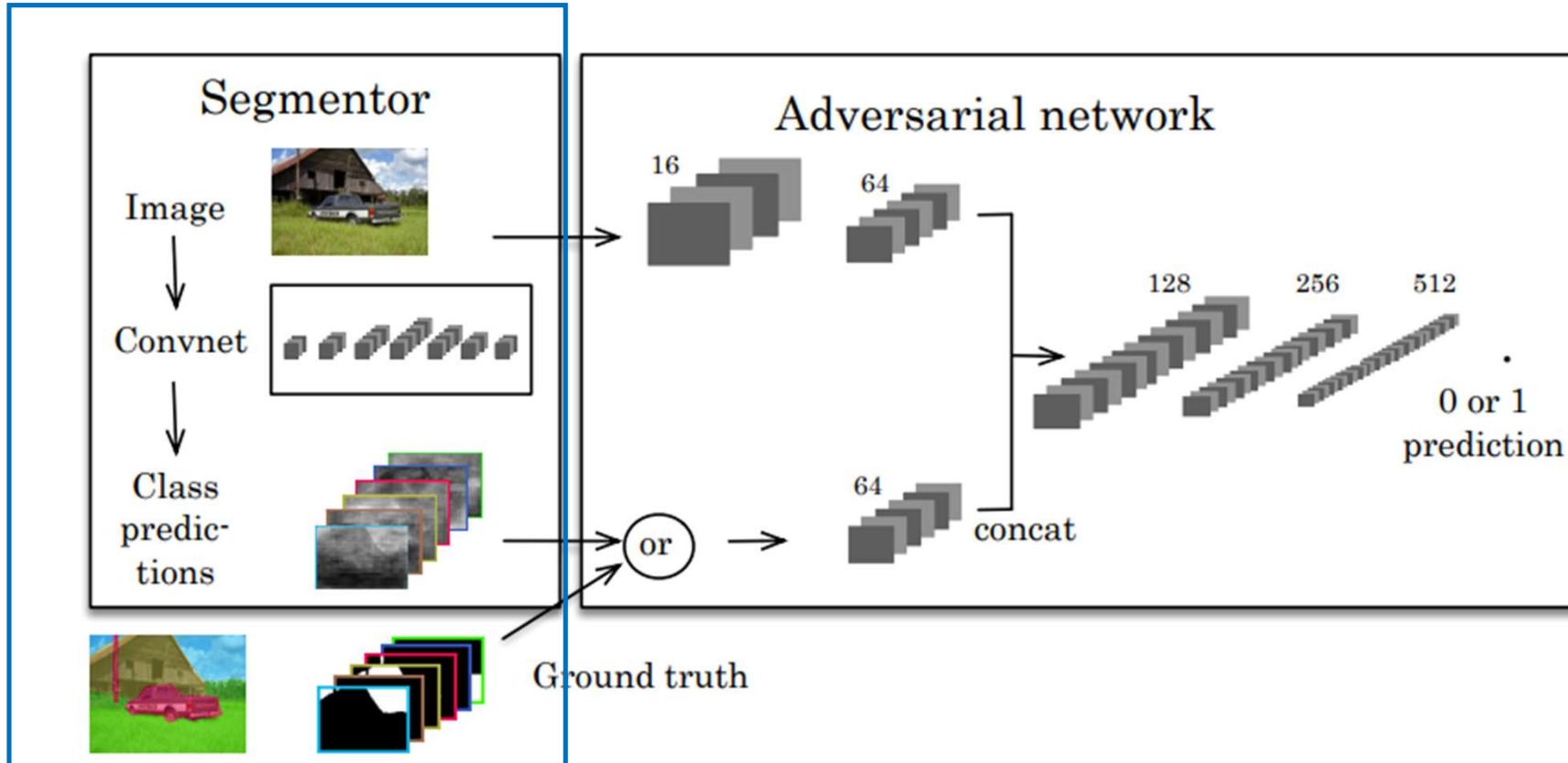
Semantic Segmentation using Adversarial Networks. NIPS 2016.

GANs for Image Segmentation



Semantic Segmentation using Adversarial Networks. NIPS 2016.

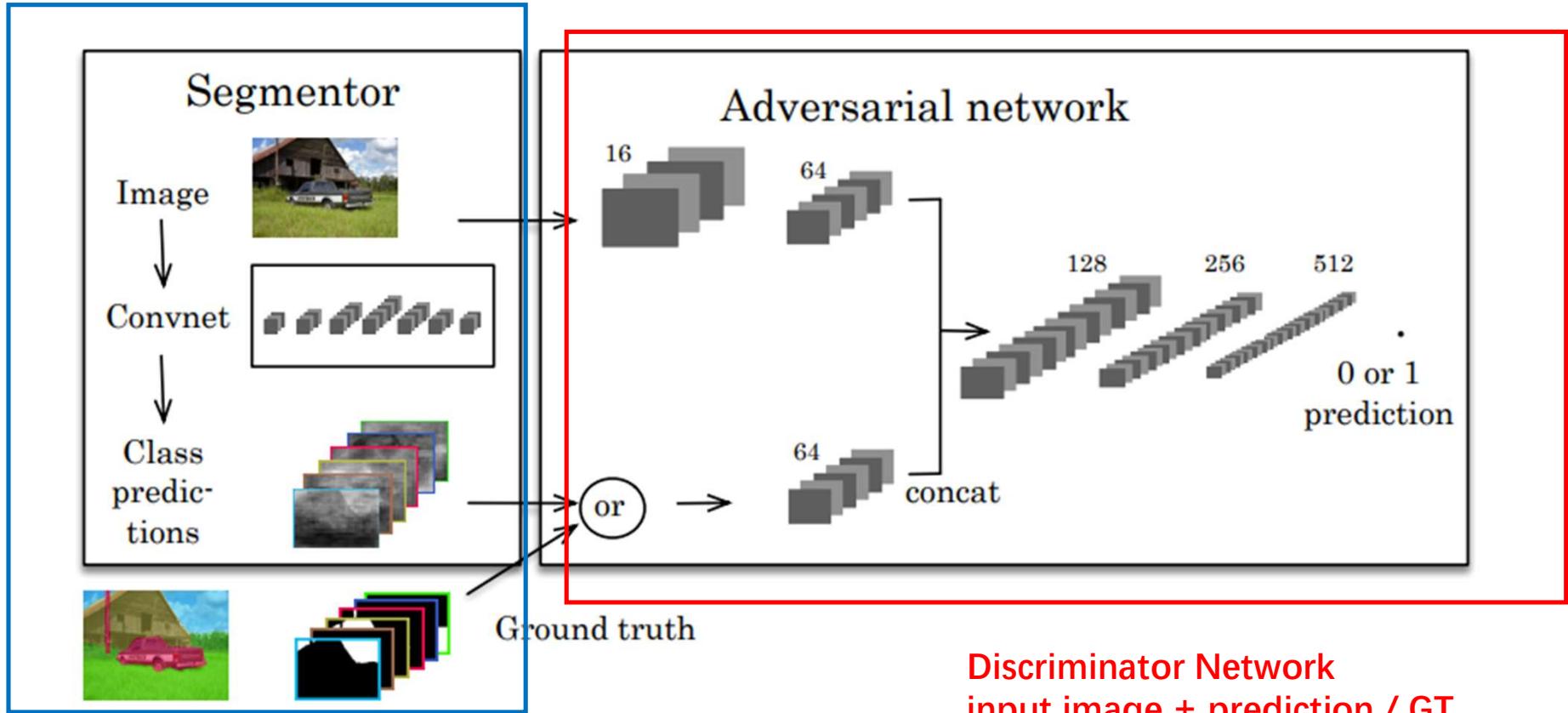
GANs for Image Segmentation



Classic Image
Segmentation
with CNN

Semantic Segmentation using Adversarial Networks. NIPS 2016.

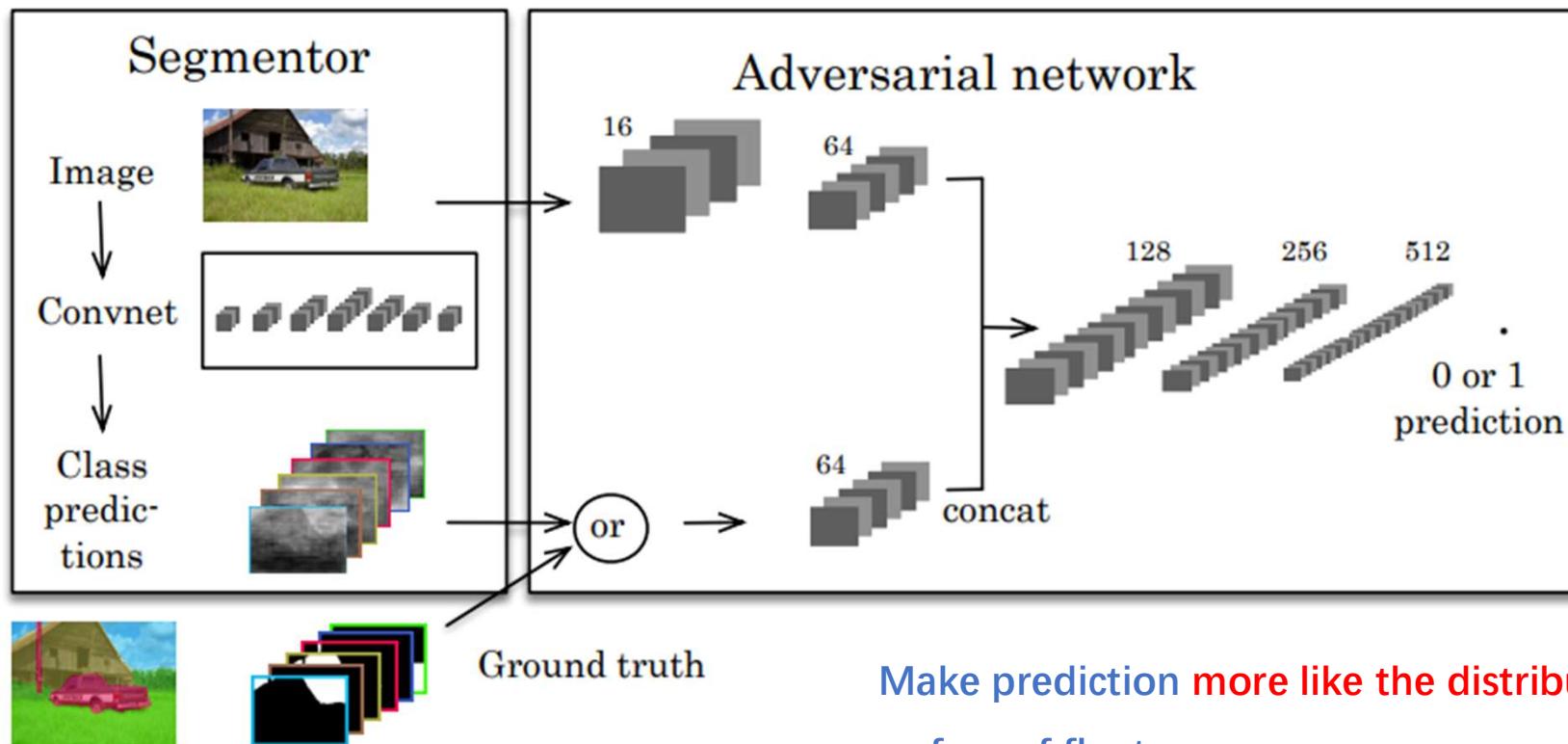
GANs for Image Segmentation



Classic Image
Segmentation
with CNN

Semantic Segmentation using Adversarial Networks. NIPS 2016.

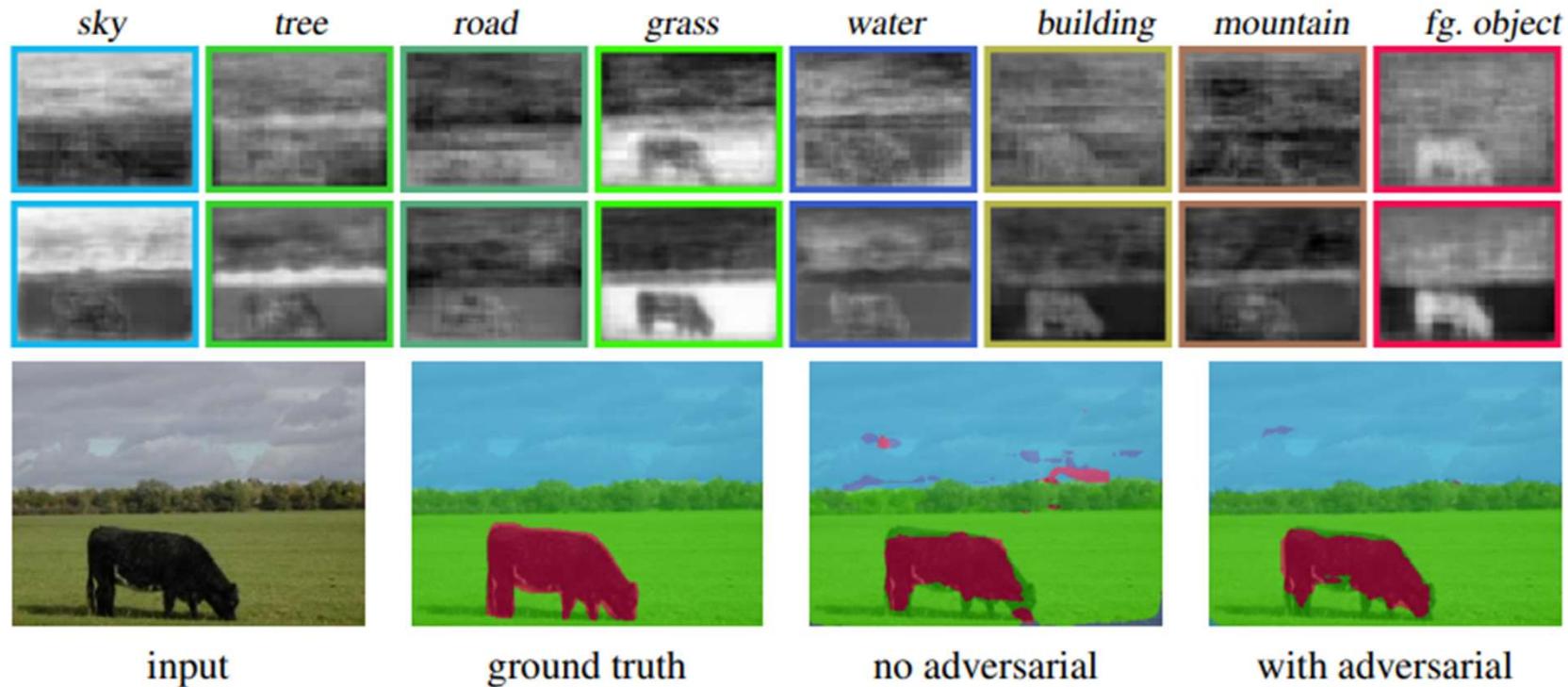
GANs for Image Segmentation



- Make prediction **more like the distribution of GT**
 - free of floaters** ...
 - preserve contiguity**

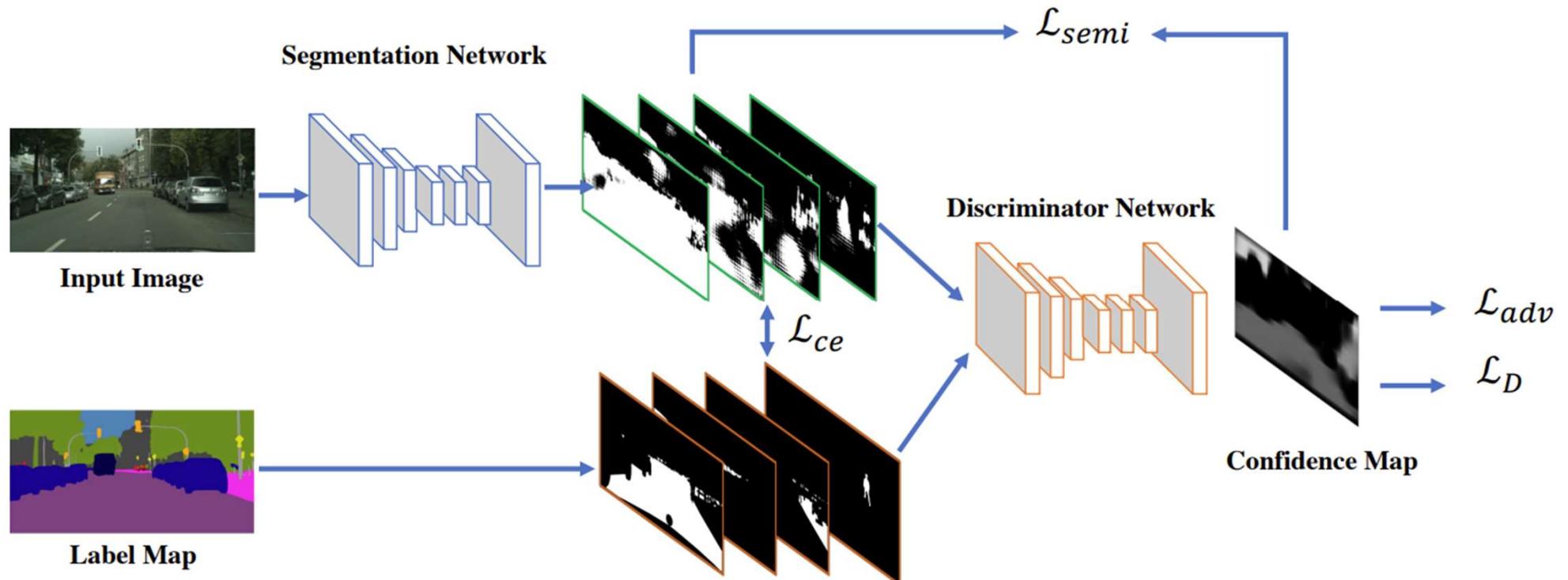
Semantic Segmentation using Adversarial Networks. NIPS 2016.

GANs for Image Segmentation



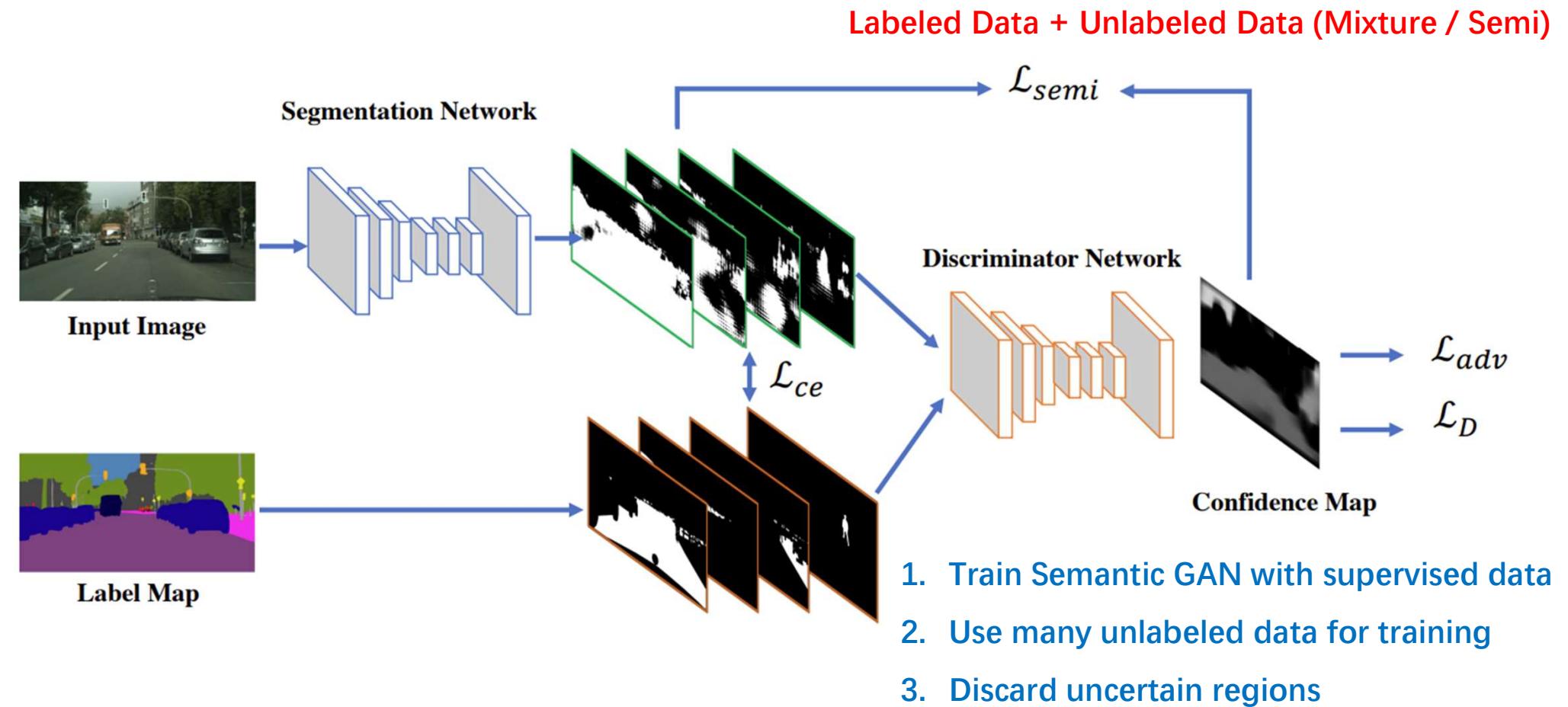
Semantic Segmentation using Adversarial Networks. NIPS 2016.

GANs for Image Segmentation



Adversarial Learning for Semi-Supervised Semantic Segmentation. BMVC 2018

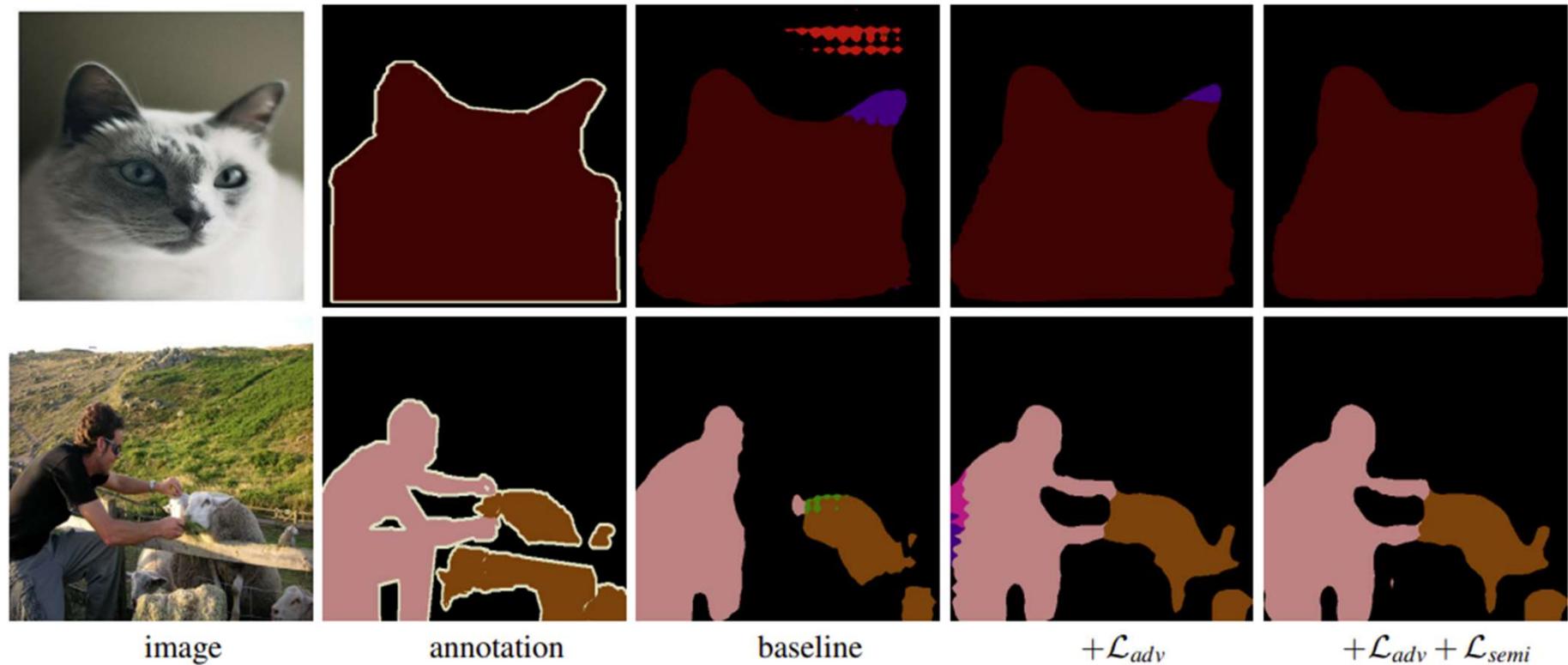
GANs for Image Segmentation



Adversarial Learning for Semi-Supervised Semantic Segmentation. BMVC 2018

GANs for Image Segmentation

Labeled Data + Unlabeled Data (Mixture / Semi)



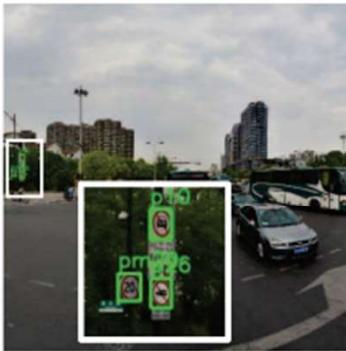
Adversarial Learning for Semi-Supervised Semantic Segmentation. BMVC 2018

GANs for Object Detection

Zhu et al. [45]



Ours



Zhu et al. [45]



Ours



Zhu et al. [45]

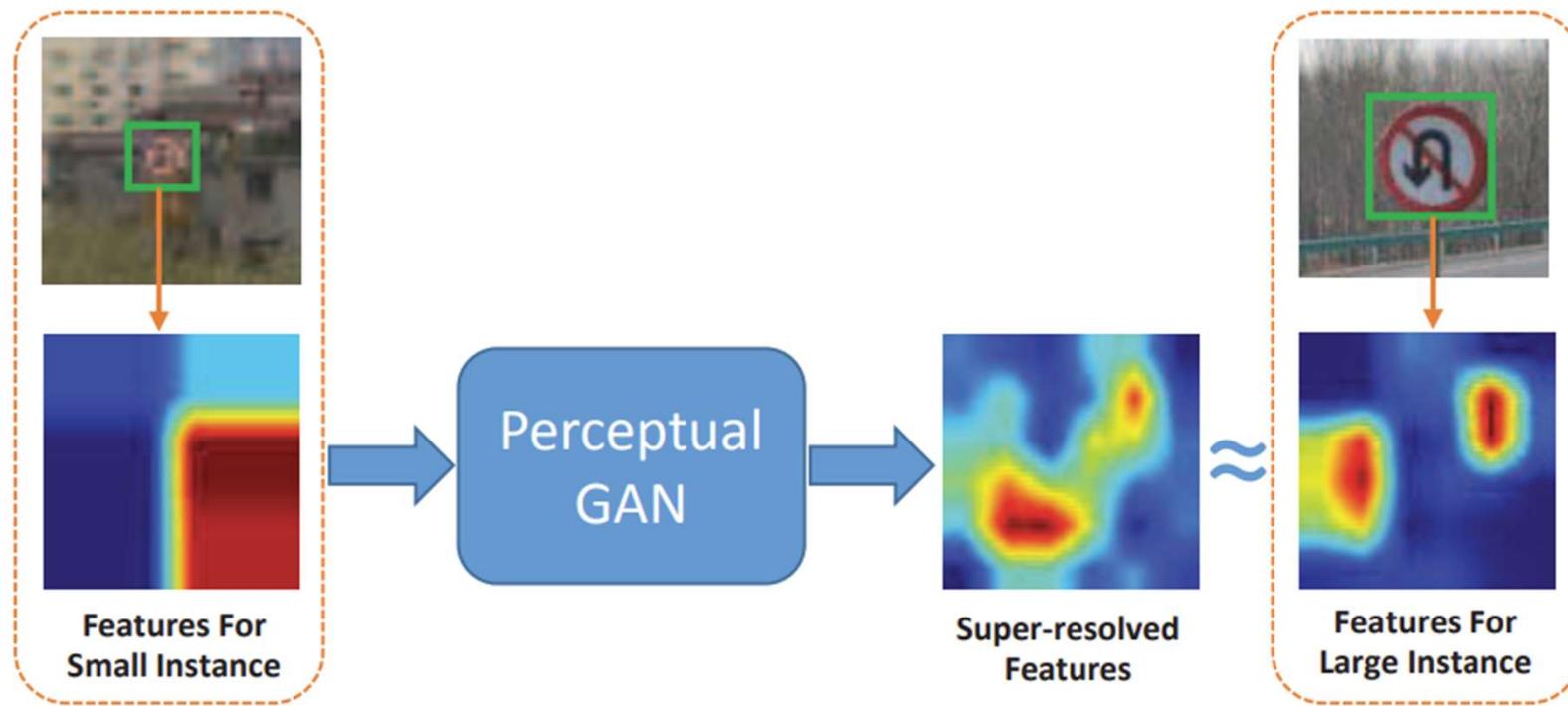


Ours



Perceptual GAN for Small Object Detection, CVPR 2017

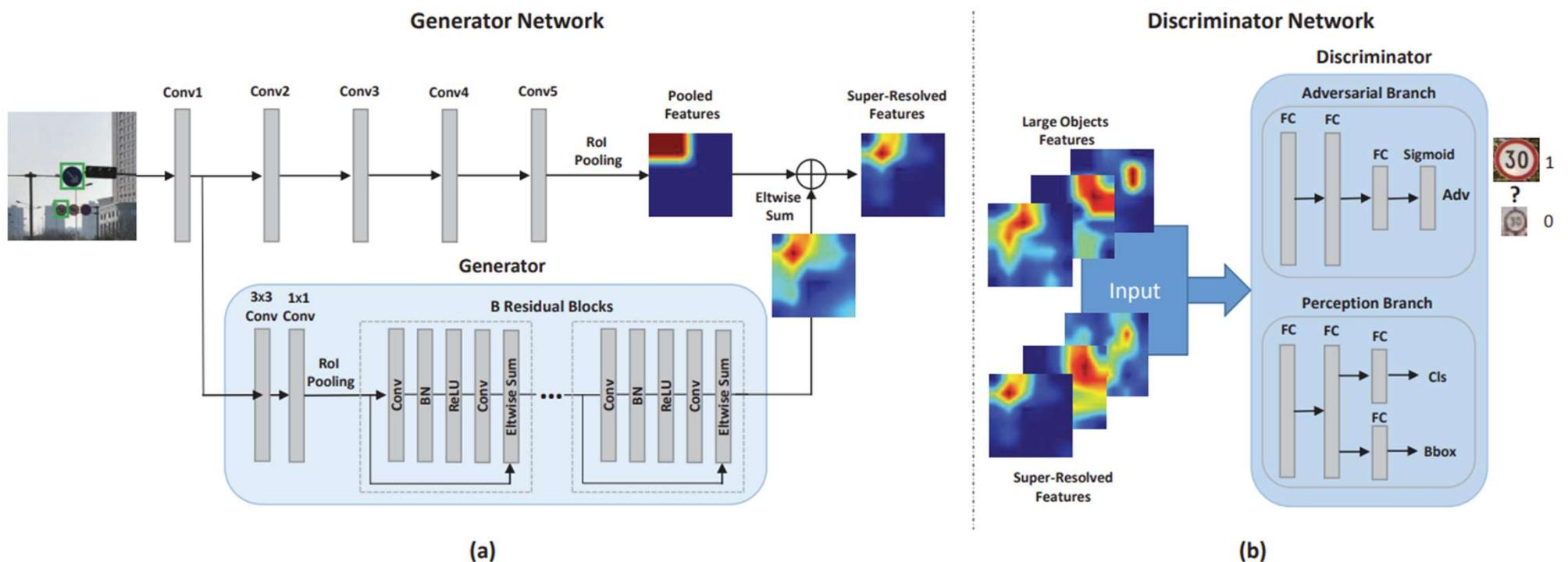
GANs for Object Detection



Using GAN to close the gap between large and small objects

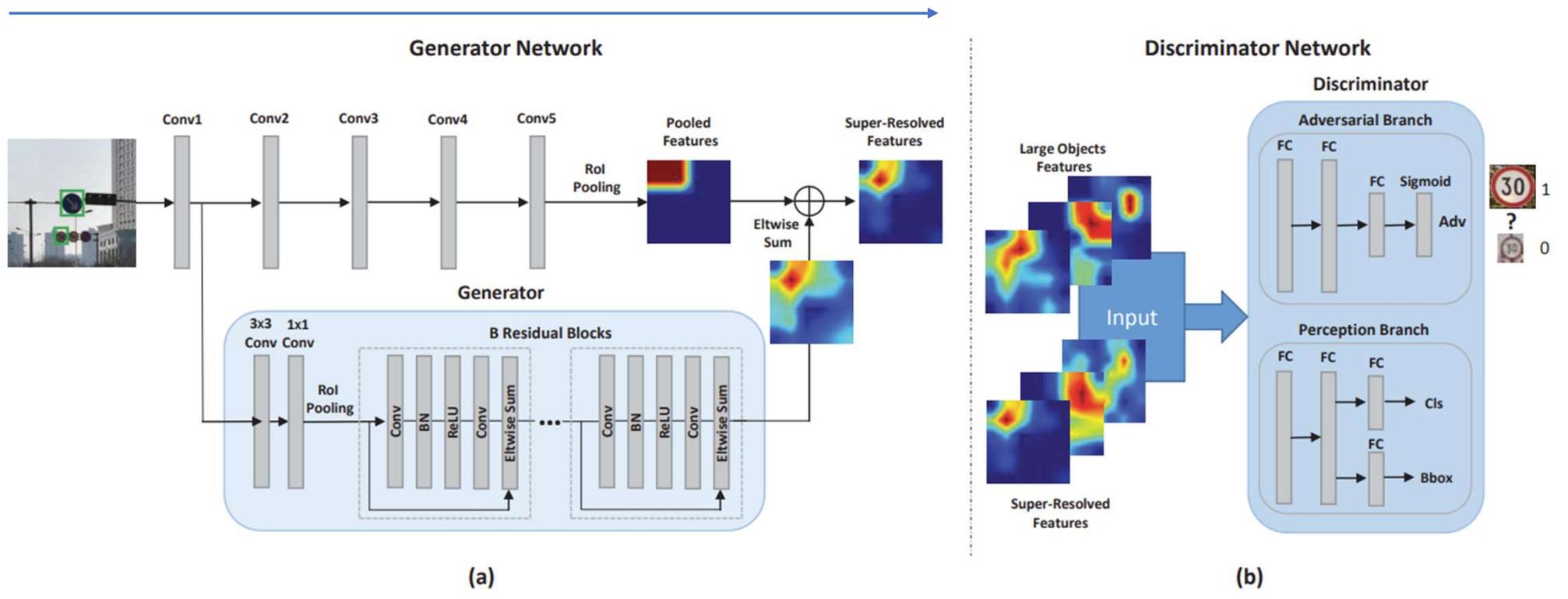
Perceptual GAN for Small Object Detection, CVPR 2017

GANs for Object Detection



Perceptual GAN for Small Object Detection, CVPR 2017

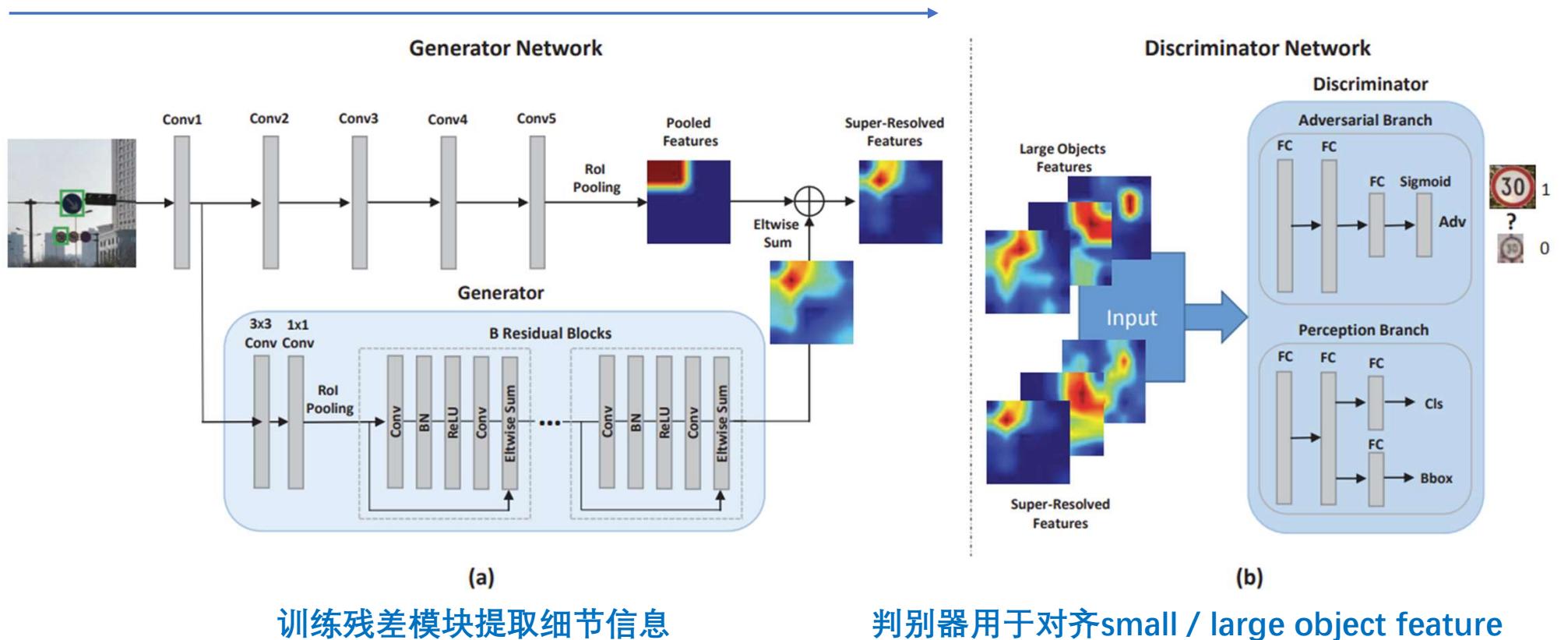
GANs for Object Detection



训练残差模块提取细节信息

Perceptual GAN for Small Object Detection, CVPR 2017

GANs for Object Detection



(a)

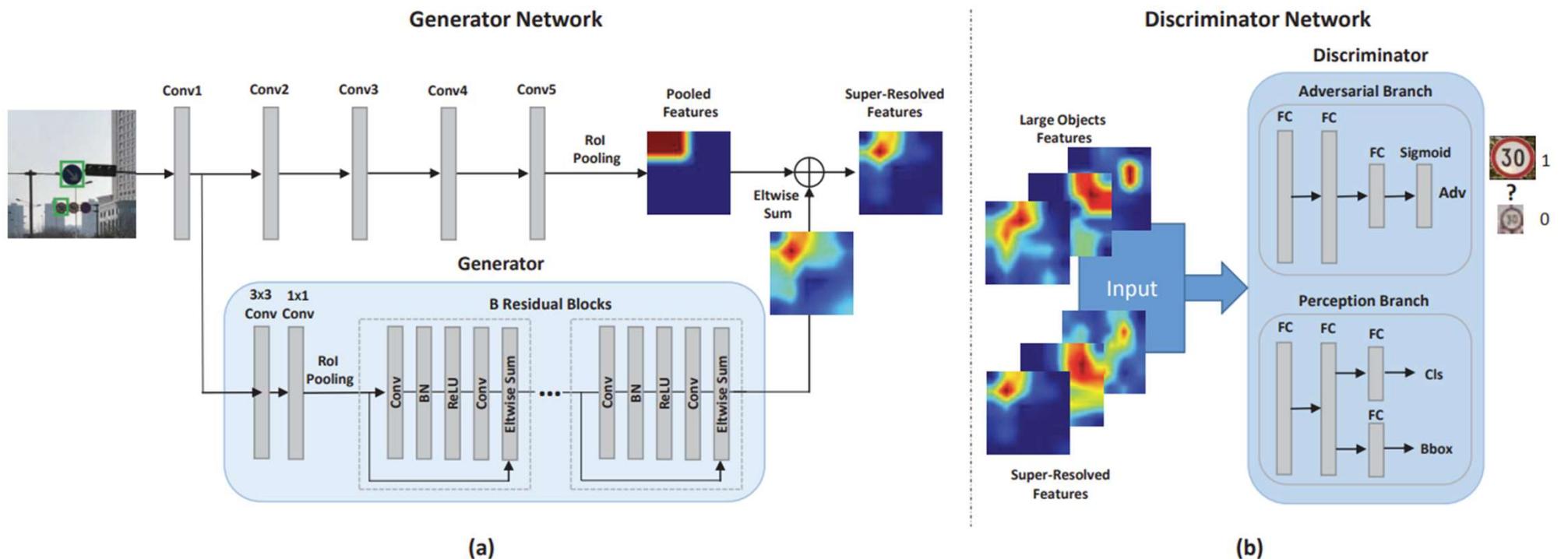
训练残差模块提取细节信息

(b)

判别器用于对齐small / large object feature

Perceptual GAN for Small Object Detection, CVPR 2017

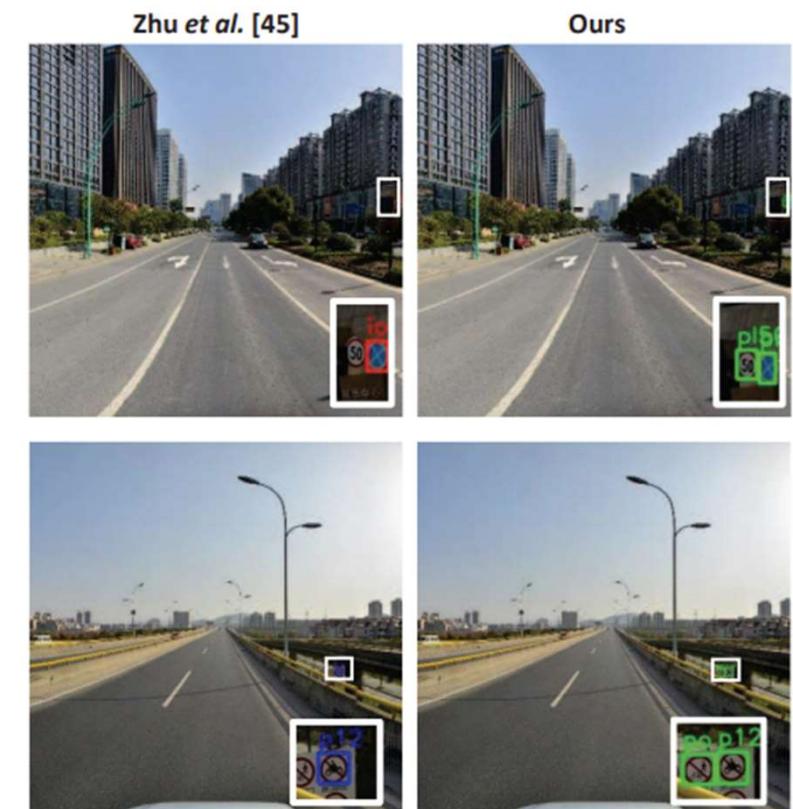
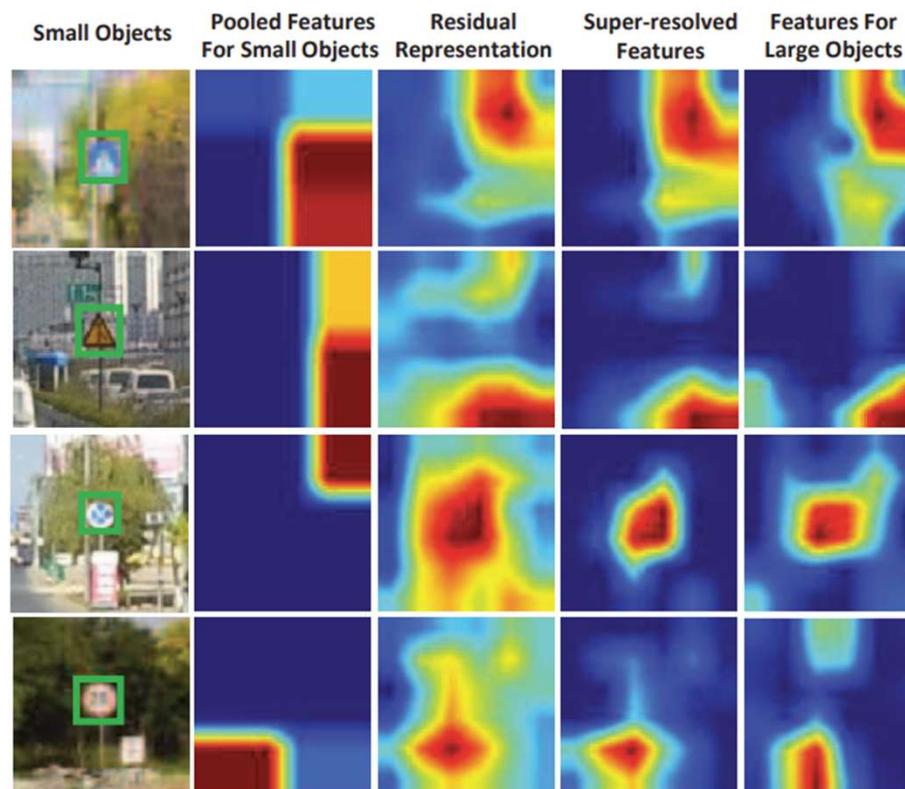
GANs for Object Detection



Make feature of small objects be like the feature distribution of large objects

Perceptual GAN for Small Object Detection, CVPR 2017

GANs for Object Detection

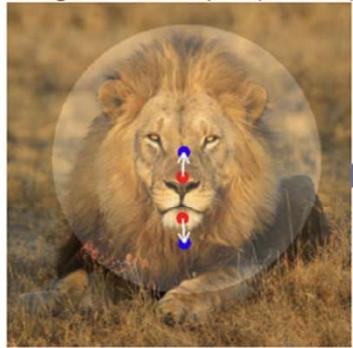


Make feature of small objects be like the feature distribution of large objects

Perceptual GAN for Small Object Detection, CVPR 2017

GANs for Geometric Image Editing

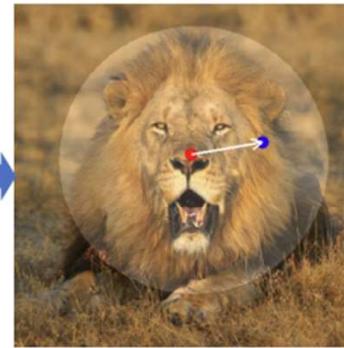
Image + User input (1st Edit)



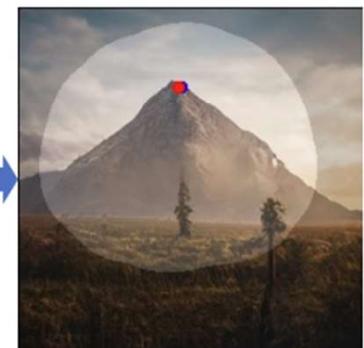
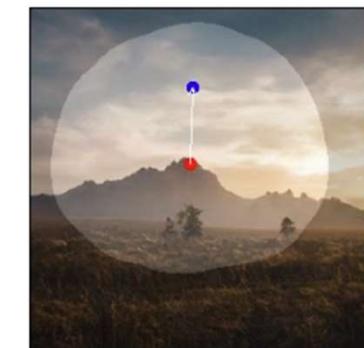
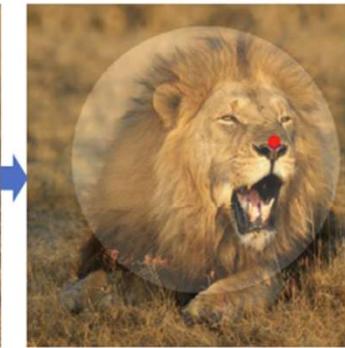
Result



2nd Edit

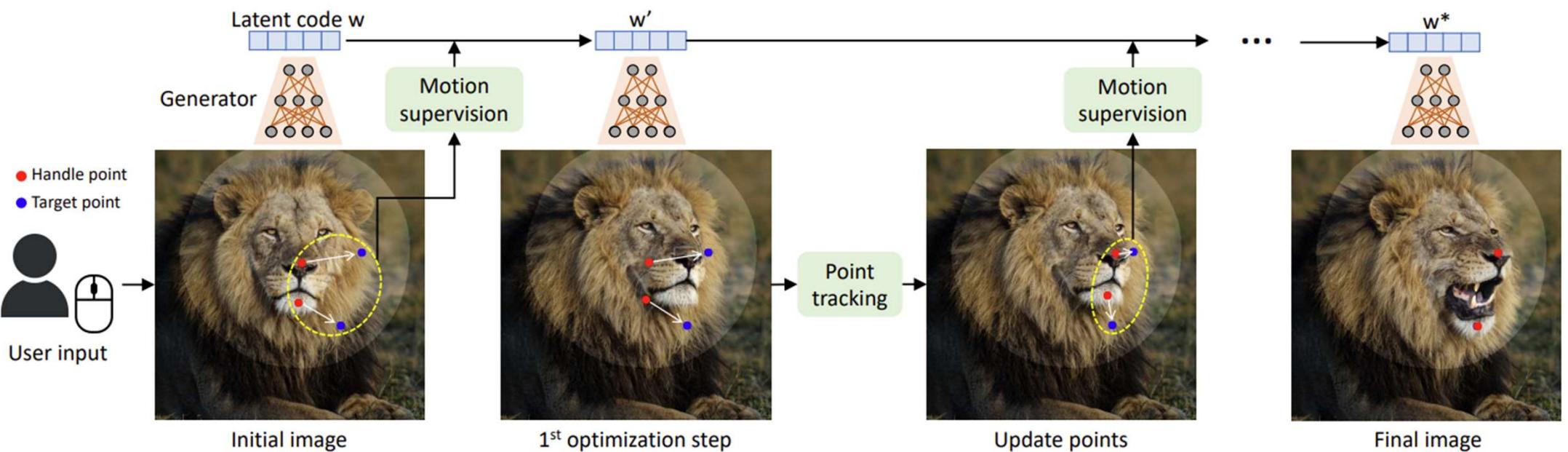


Result



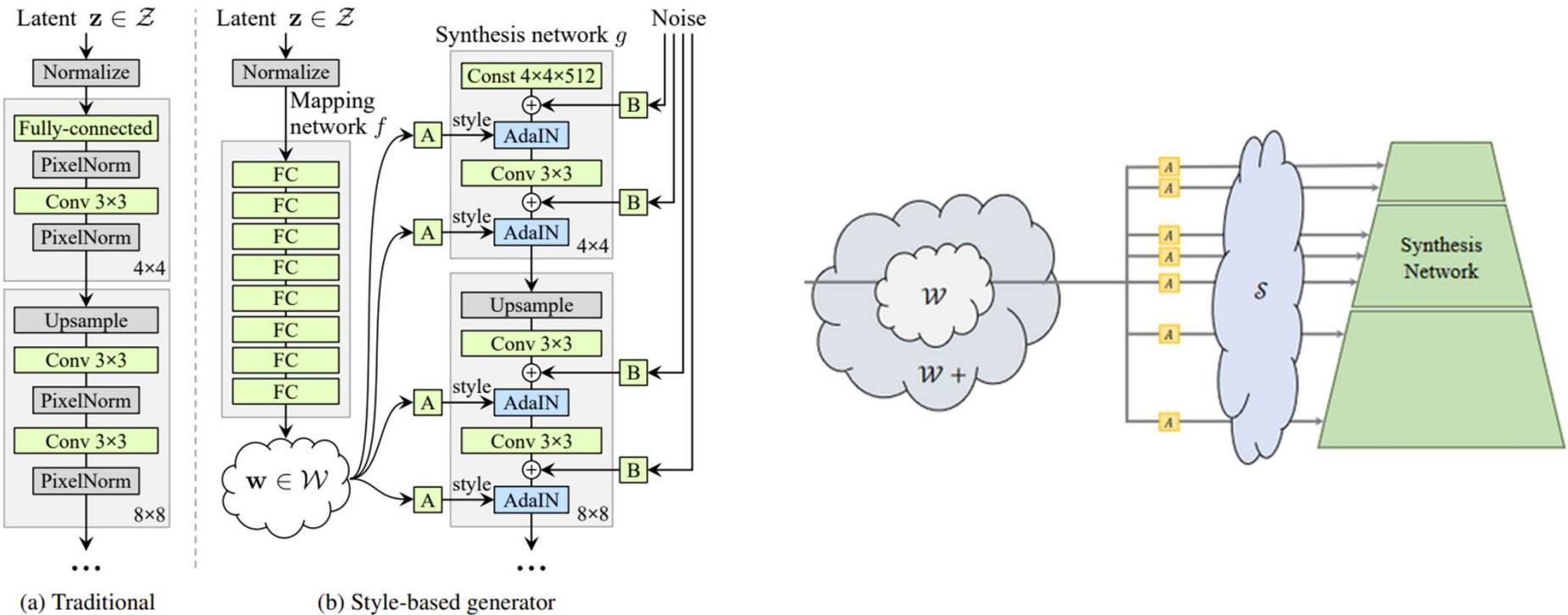
Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Geometric Image Editing



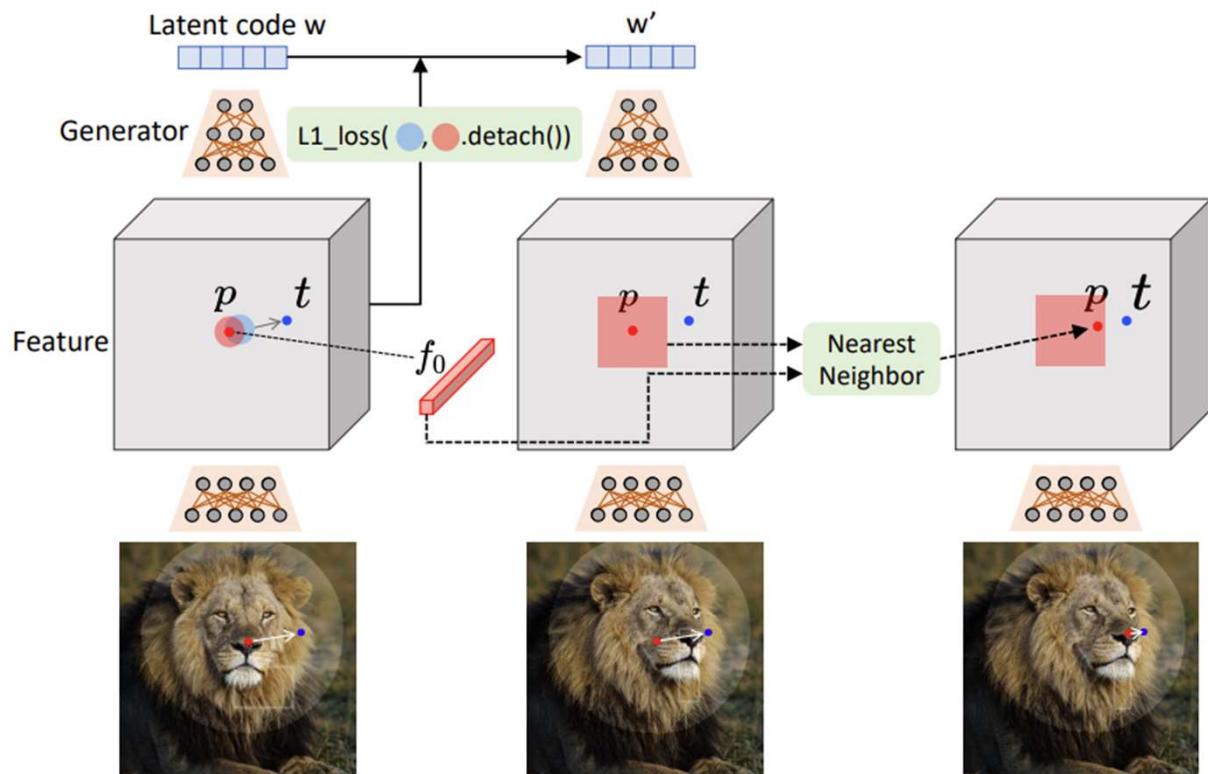
Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Geometric Image Editing



Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

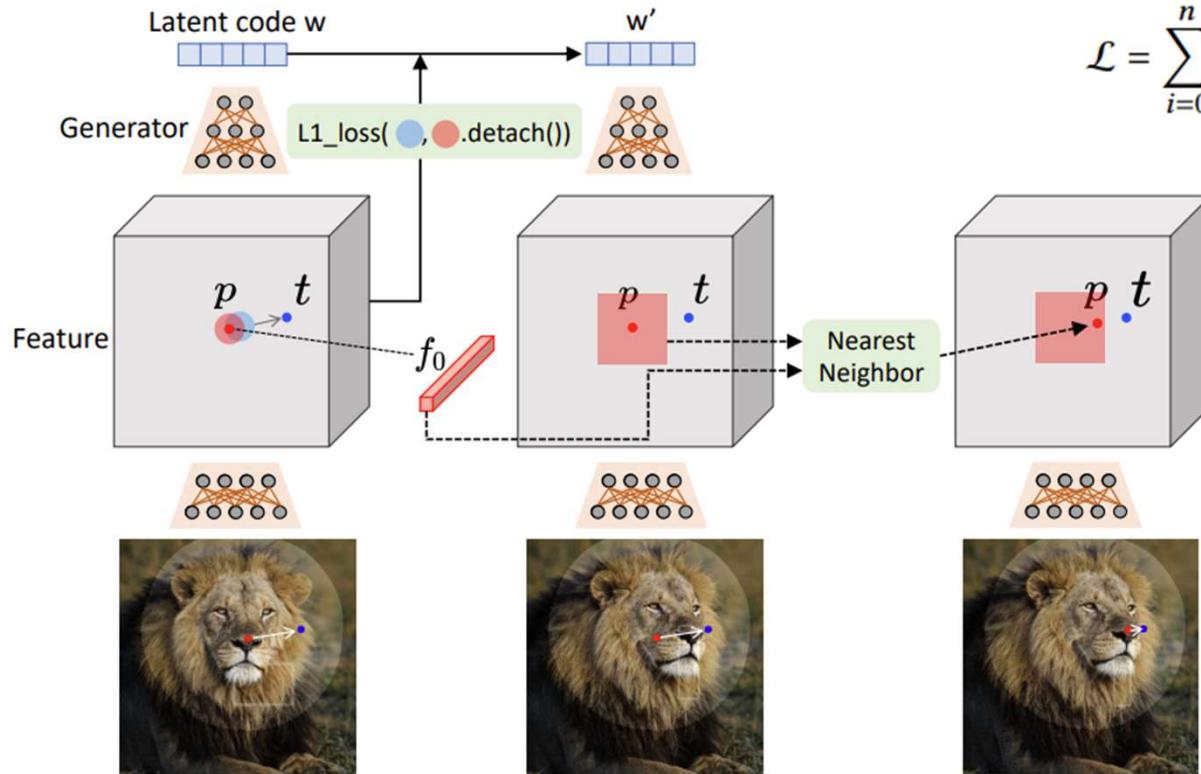
GANs for Geometric Image Editing



Optimize $w \rightarrow w'$ in W space

Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Geometric Image Editing



$$\mathcal{L} = \sum_{i=0}^n \sum_{q_i \in \Omega_1(p_i, r_1)} \|\mathbf{F}(q_i) - \mathbf{F}(q_i + d_i)\|_1 + \lambda \|\mathbf{F} - \mathbf{F}_0\|_1 \cdot (1 - \mathbf{M}) \quad (1)$$

$$d_i = \frac{t_i - p_i}{\|t_i - p_i\|_2}$$

Since $\mathbf{F}(q_i)$ detached

The gradient will not back-propagate through it

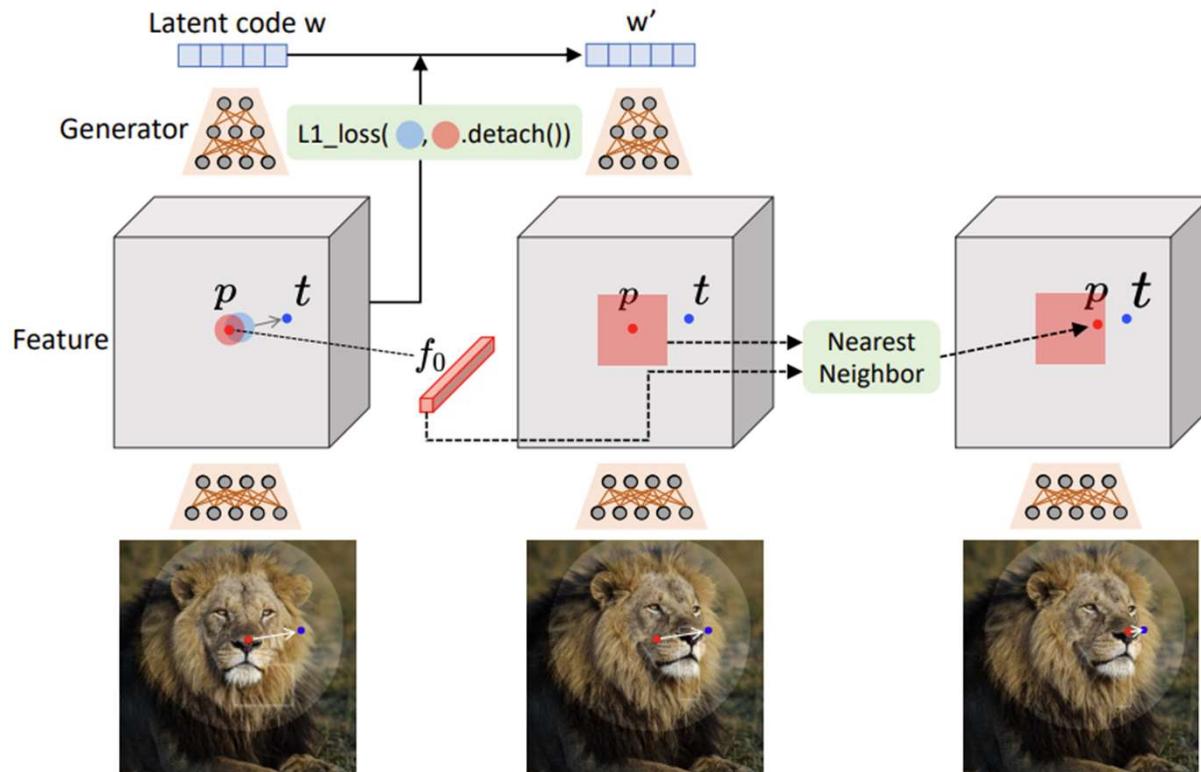
Making only one-direction move

i.e. $p_i \rightarrow p_i + d_i$

Optimize $w \rightarrow w'$ in W space

Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Geometric Image Editing



$$f_i = \mathbf{F}_0(\mathbf{p}_i)$$

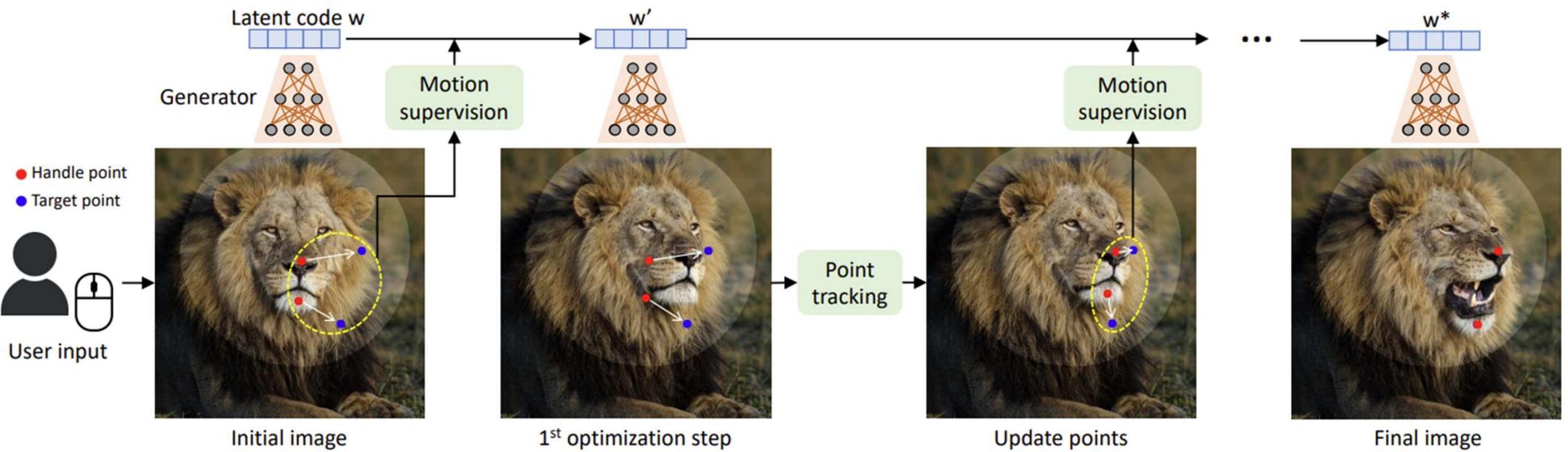
$$\mathbf{p}_i := \arg \min_{\mathbf{q}_i \in \Omega_2(\mathbf{p}_i, r_2)} \|\mathbf{F}'(\mathbf{q}_i) - f_i\|_1$$

Precisely tracking point and
Move to the target Iteratively

Optimize $w \rightarrow w'$ in W space

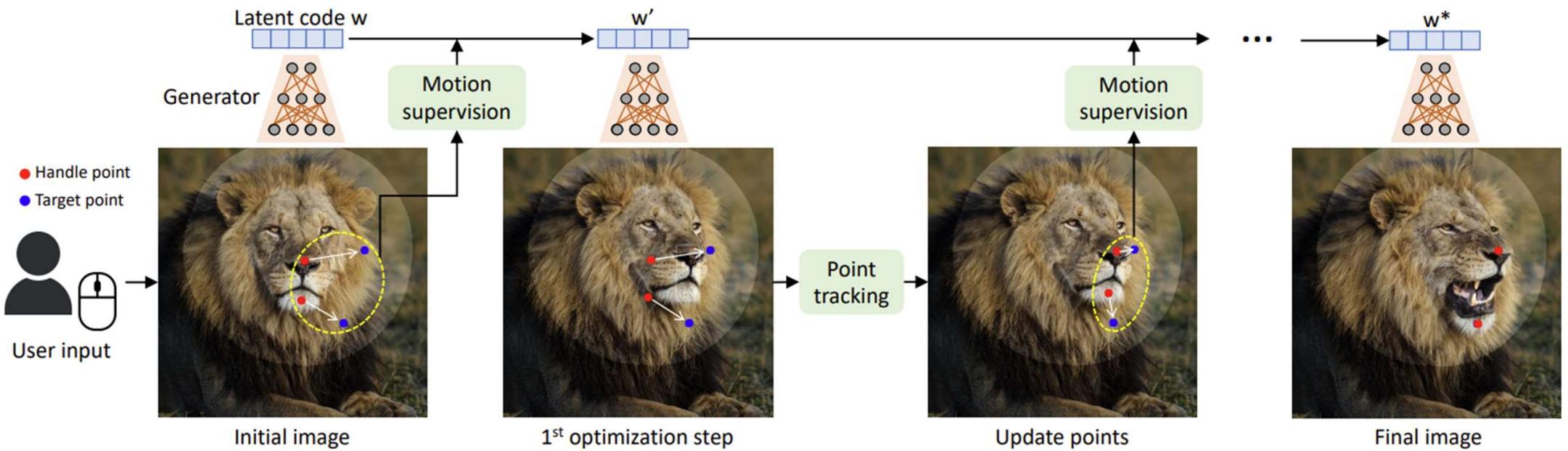
Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Geometric Image Editing



Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Geometric Image Editing



Not explicitly make output be like real, but utilize the latent code space of pretrained StyleGAN
for Precise and Realistic Editing

Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Geometric Image Editing



Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold, SIGGRAPH 2023.

GANs for Image Texture / Color Editing

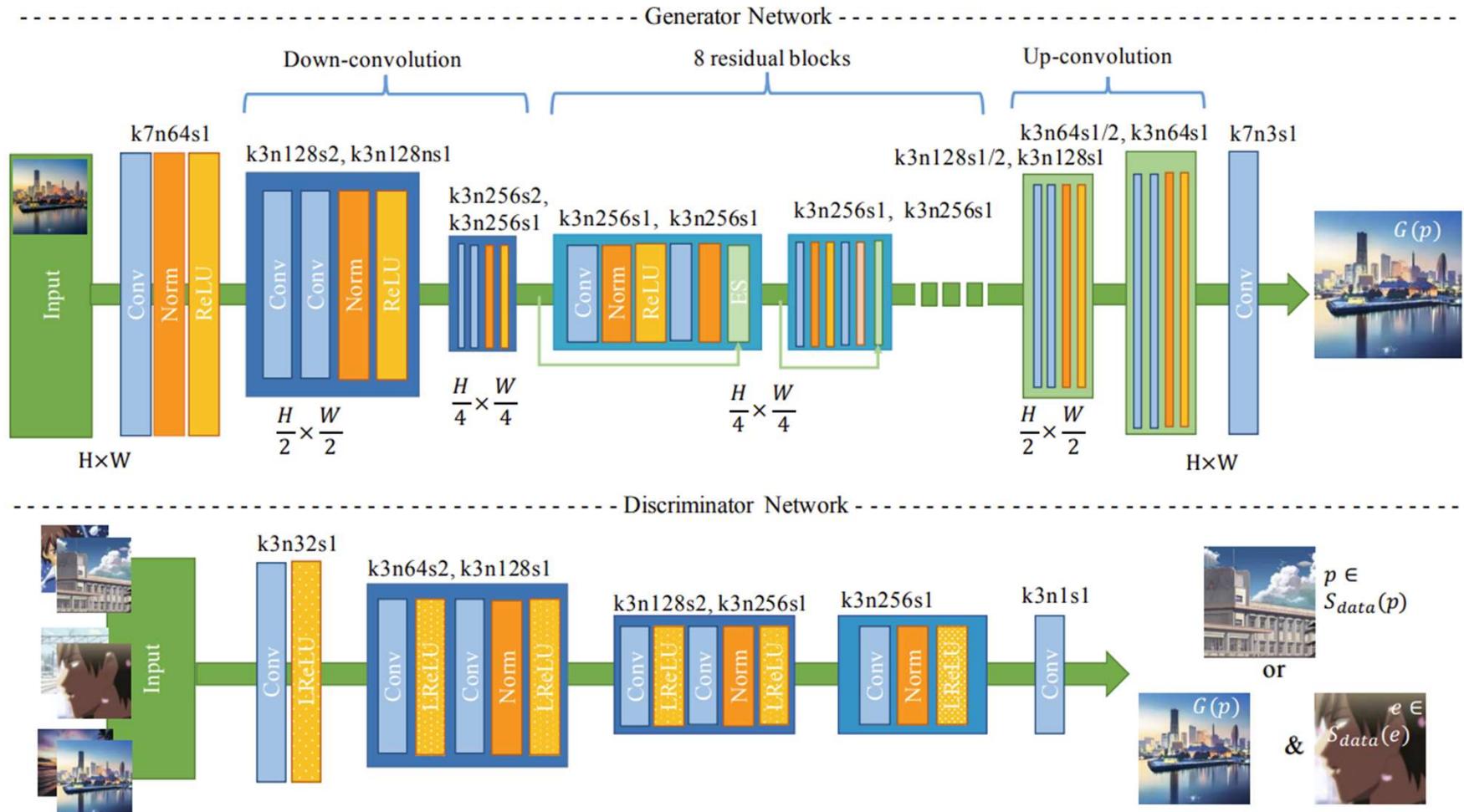


GANs for Image Texture / Color Editing



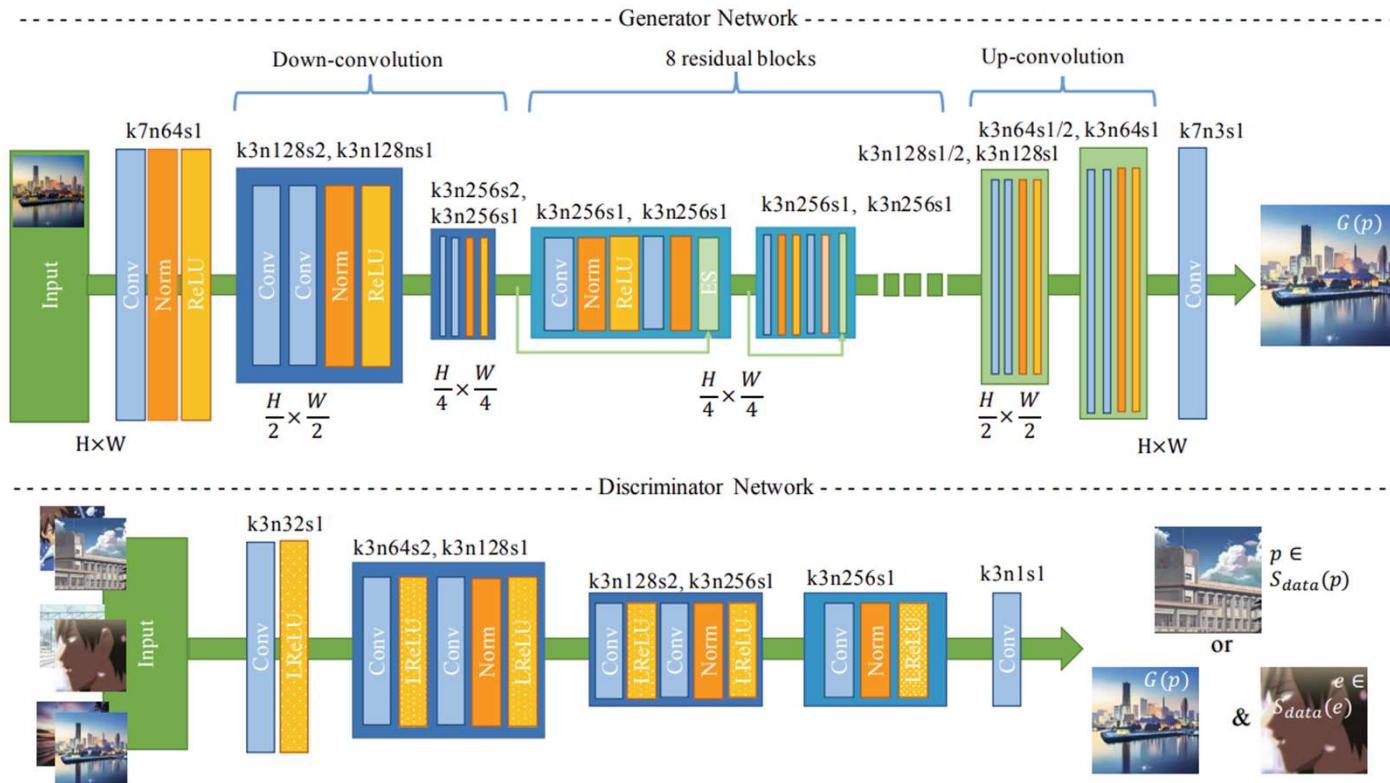
Generative Adversarial Networks for photo to Hayao Miyazaki style cartoons, 2020.

GANs for Image Texture / Color Editing



CartoonGAN: Generative Adversarial Networks for Photo Cartoonization, CVPR 2018.

GANs for Image Texture / Color Editing



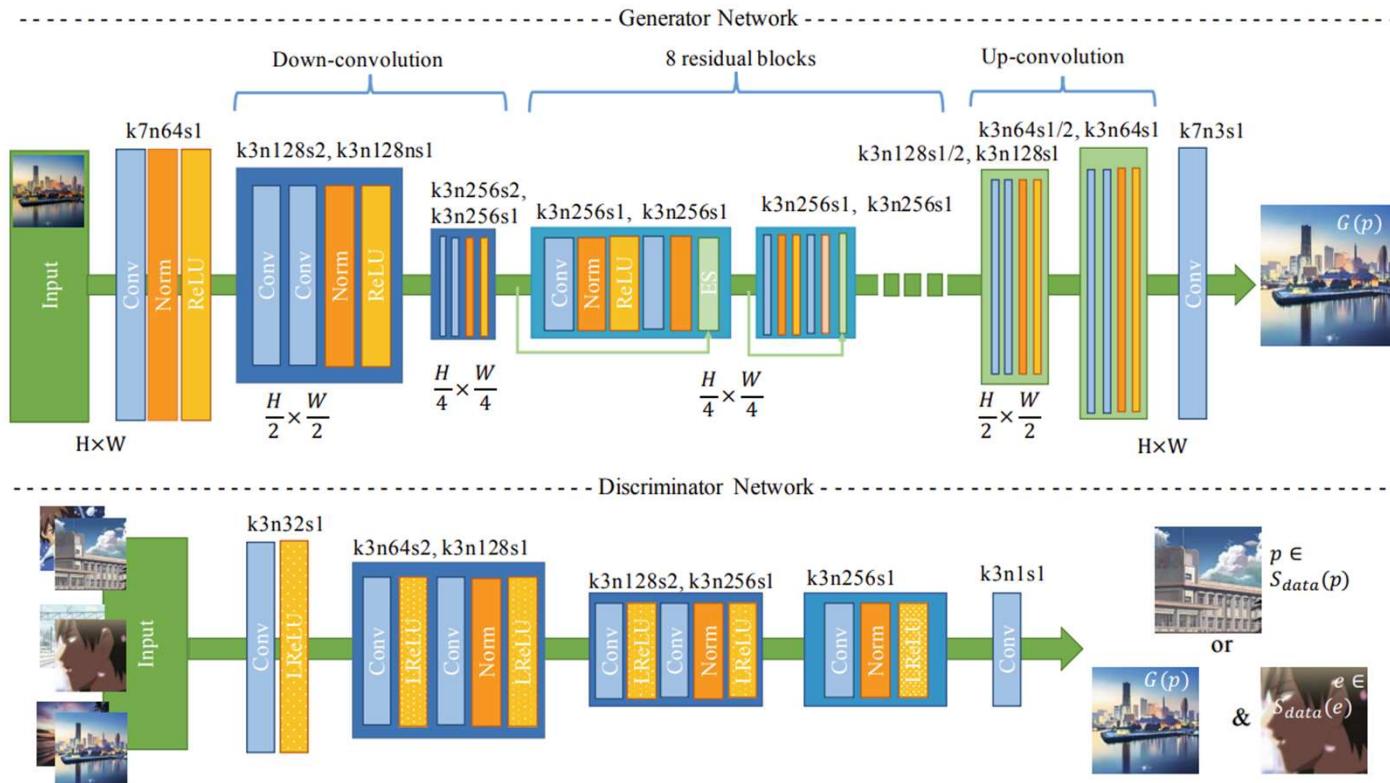
$$\mathcal{L}(G, D) = \mathcal{L}_{adv}(G, D) + \omega \mathcal{L}_{con}(G, D)$$

$$\begin{aligned} \mathcal{L}_{adv}(G, D) = & \mathbb{E}_{c_i \sim S_{data}(c)} [\log D(c_i)] \\ & + \mathbb{E}_{e_j \sim S_{data}(e)} [\log(1 - D(e_j))] \\ & + \mathbb{E}_{p_k \sim S_{data}(p)} [\log(1 - D(G(p_k)))] \end{aligned}$$



CartoonGAN: Generative Adversarial Networks for Photo Cartoonization, CVPR 2018.

GANs for Image Texture / Color Editing



$$\mathcal{L}(G, D) = \mathcal{L}_{adv}(G, D) + \omega \mathcal{L}_{con}(G, D)$$

$$\mathcal{L}_{con}(G, D) =$$

$$\mathbb{E}_{p_i \sim S_{data}(p)} [||VGG_l(G(p_i)) - VGG_l(p_i)||_1]$$

Make output be like cartoon images while keeping original structure

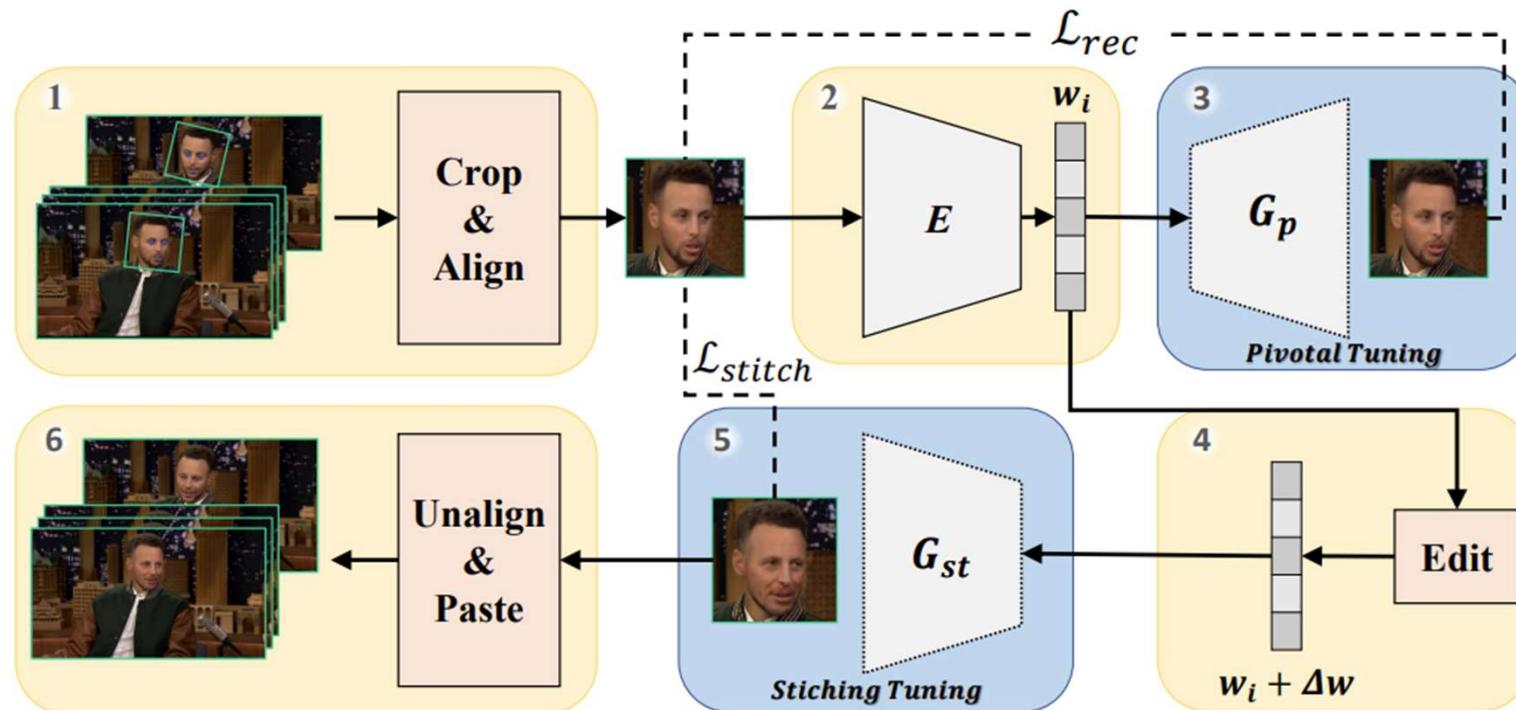
CartoonGAN: Generative Adversarial Networks for Photo Cartoonization, CVPR 2018.

GANs for Video Geo/Color Editing



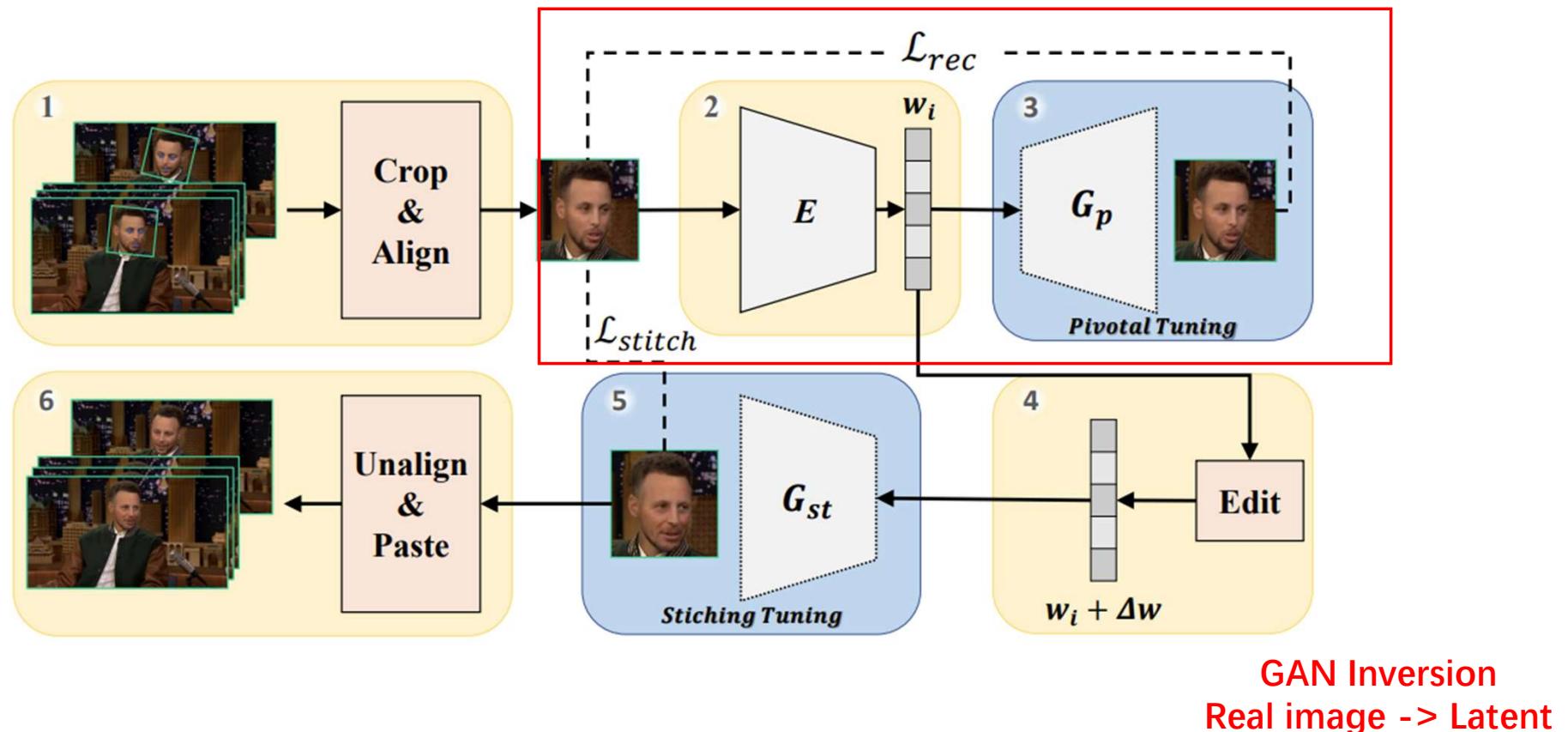
Stitch it in Time: GAN-Based Facial Editing of Real Videos, SIGGRAPH ASIA 2022.

GANs for Video Geo/Color Editing



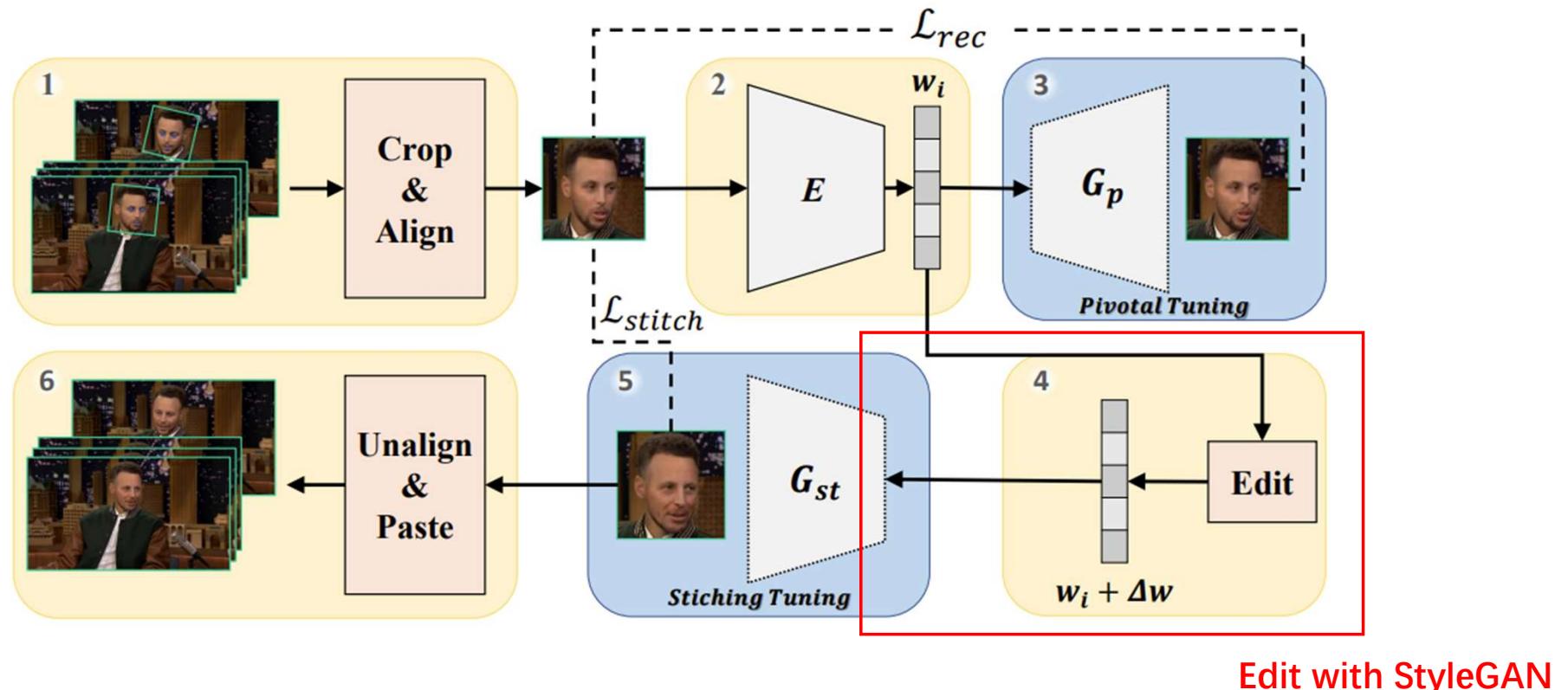
Stitch it in Time: GAN-Based Facial Editing of Real Videos, SIGGRAPH ASIA 2022.

GANs for Video Geo/Color Editing



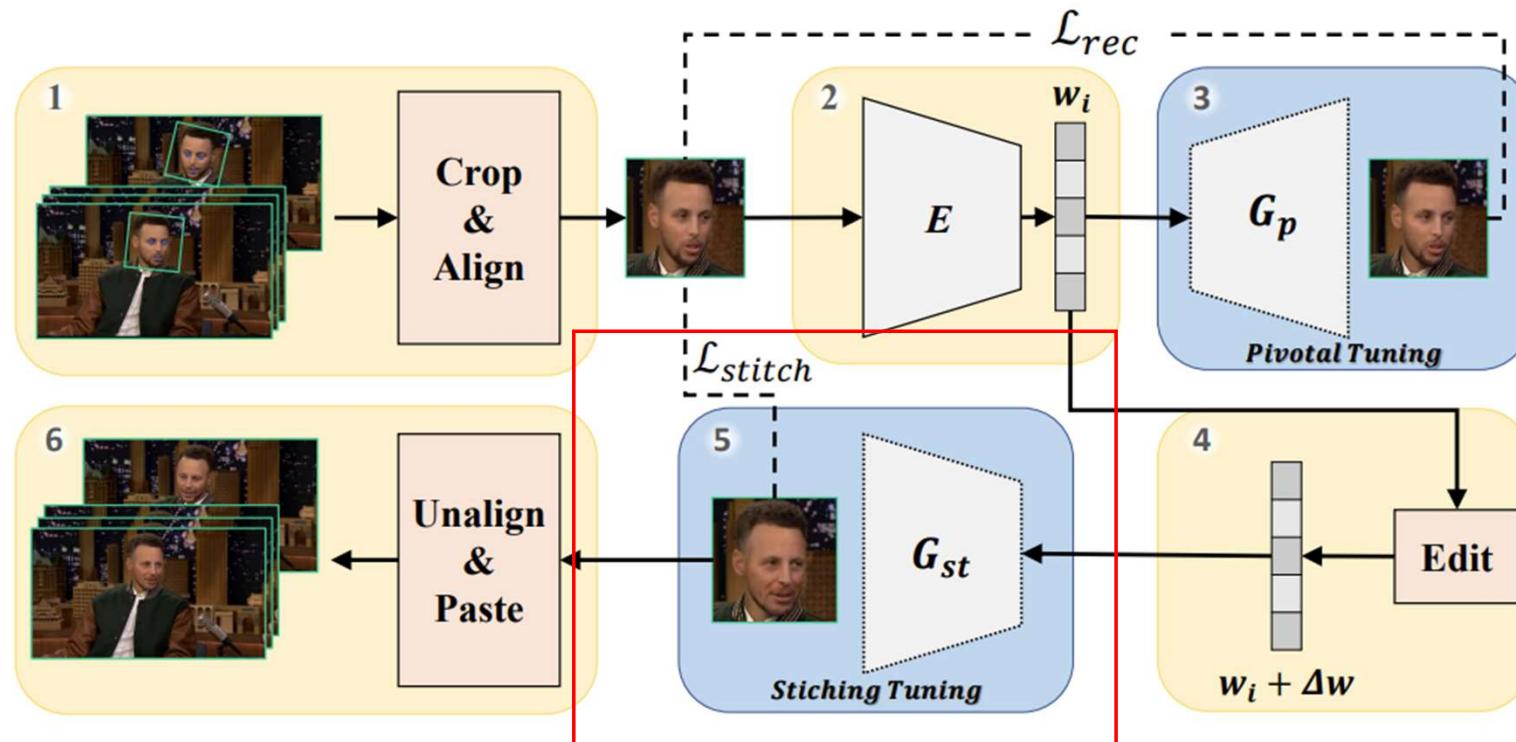
Stitch it in Time: GAN-Based Facial Editing of Real Videos, SIGGRAPH ASIA 2022.

GANs for Video Geo/Color Editing



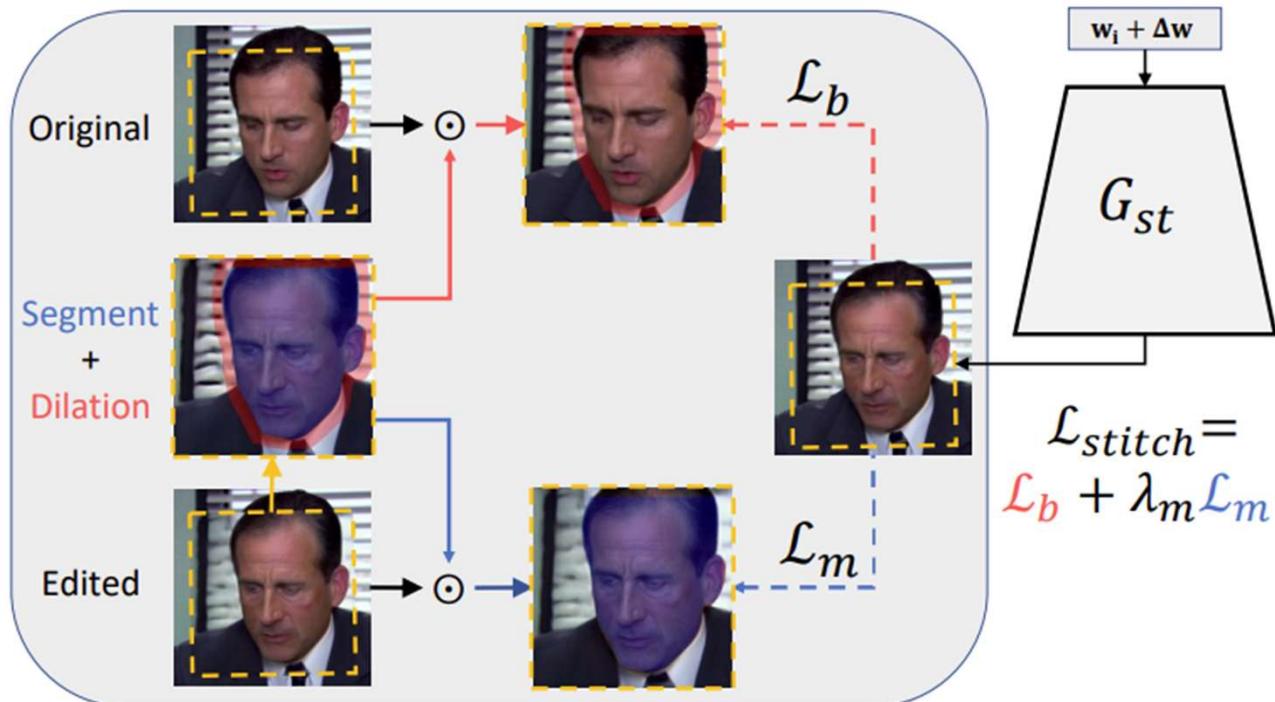
Stitch it in Time: GAN-Based Facial Editing of Real Videos, SIGGRAPH ASIA 2022.

GANs for Video Geo/Color Editing



Stitch it in Time: GAN-Based Facial Editing of Real Videos, SIGGRAPH ASIA 2022.

GANs for Video Geo/Color Editing



Blending with original boundary

Keep editing facial region

GANs for Video Geo/Color Editing



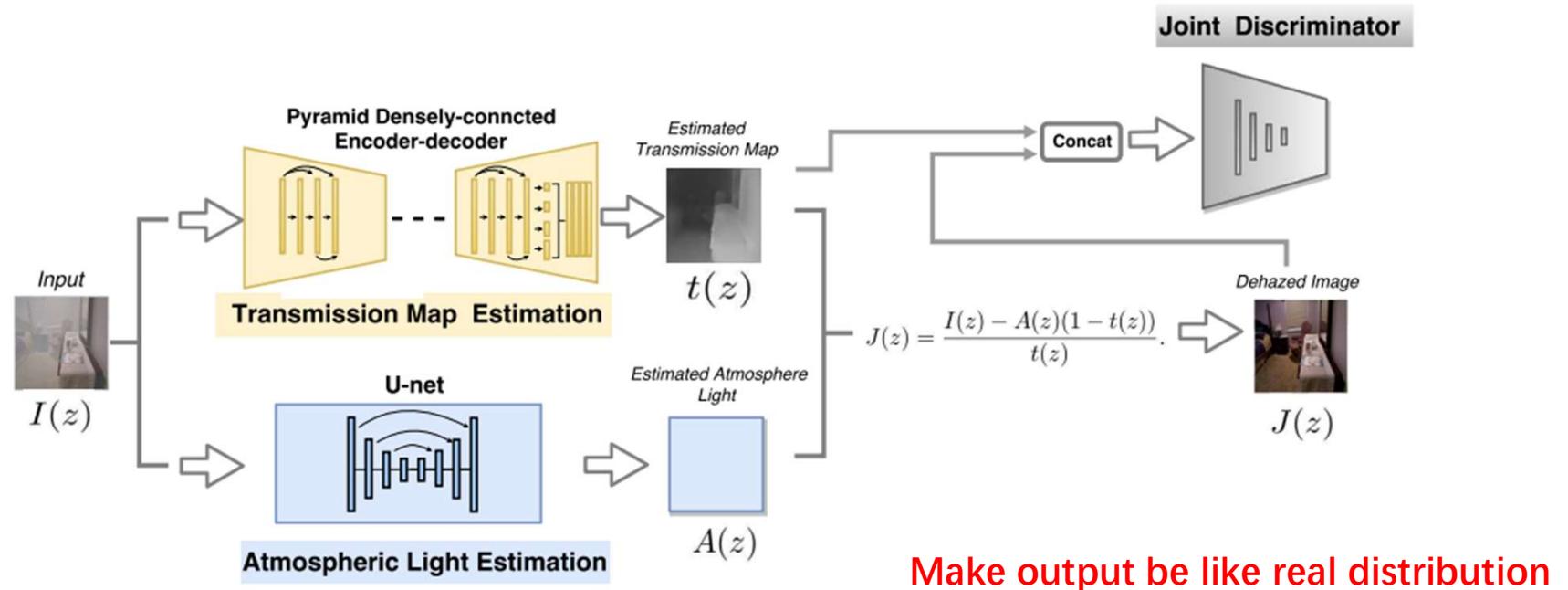
Edit in w space with StyleGAN
while keeping dynamic boundary

Stitch it in Time: GAN-Based Facial Editing of Real Videos, SIGGRAPH ASIA 2022.

Methodology

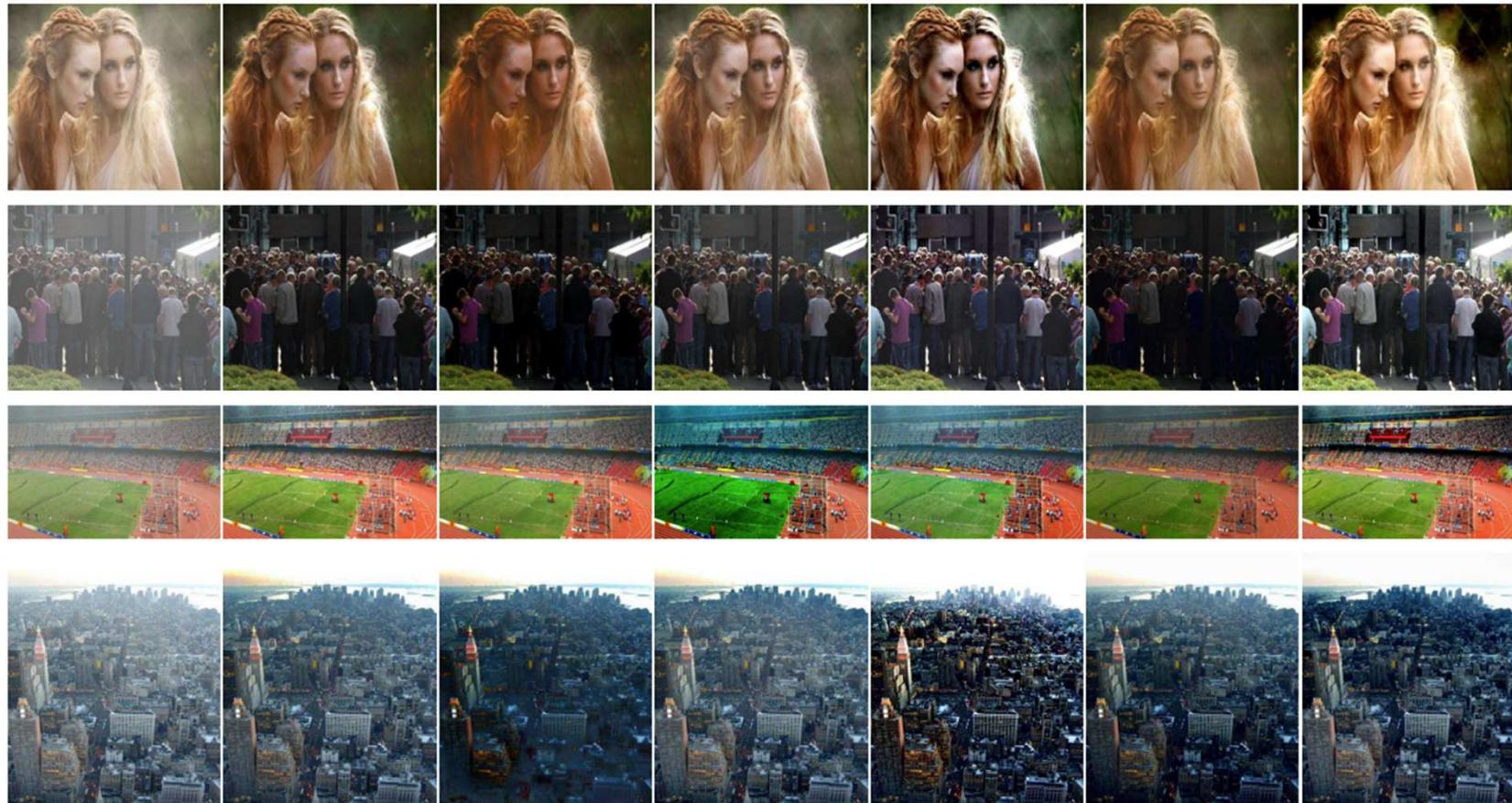
1. Make output **be like** ground truth
2. Make feature space **be like** others
3. Make output **be like** others while **keeping something**
4. Edit in **(Style)GAN latent space** (while keeping something)

Validate the Methodology: Dehaze



Densely Connected Pyramid Dehazing Network, CVPR 2018.

GAN based Dehaze Results

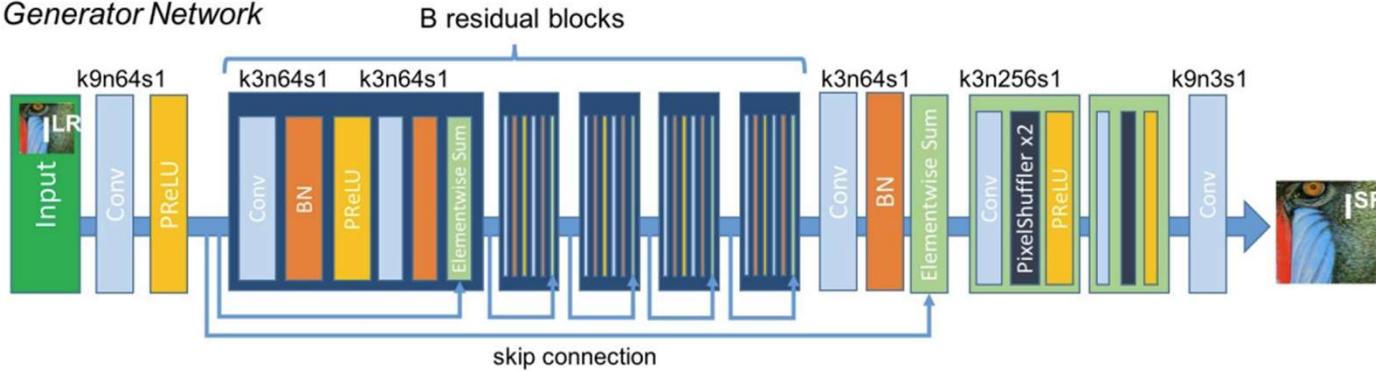


Input He. *et al.*
(CVPR'09) [13] Zhu. *et al.*
(TIP'15) [58] Ren. *et al.*
(ECCV'16)[33] Berman. *et al.*
(CVPR'16) [3, 4] Li. *et al.*
(ICCV'17) [24] DCPDN

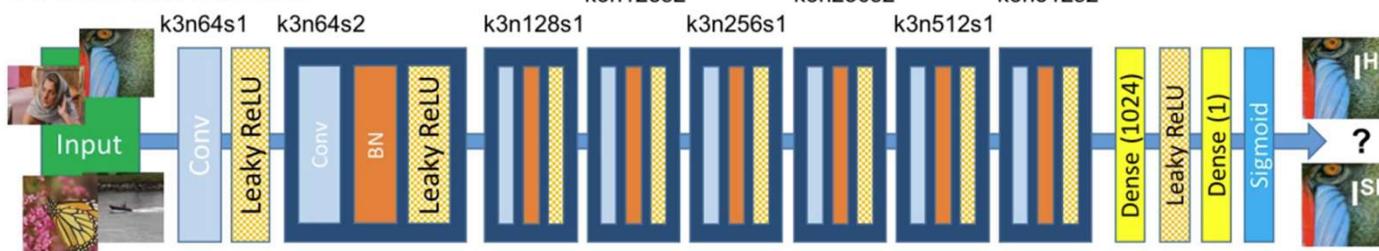
Figure 8: Dehazing results evaluated on real-world images released by the authors of previous methods.

The Methodology: Super Resolution

Generator Network



Discriminator Network



$$\min_{\theta_G} \max_{\theta_D} \mathbb{E}_{I^{HR} \sim p_{\text{train}}(I^{HR})} [\log D_{\theta_D}(I^{HR})] + \mathbb{E}_{I^{LR} \sim p_G(I^{LR})} [\log(1 - D_{\theta_D}(G_{\theta_G}(I^{LR})))]$$

2.2.1 Content loss

The pixel-wise **MSE loss** is calculated as:

$$l_{MSE}^{SR} = \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - G_{\theta_G}(I^{LR})_{x,y})^2$$

Make output be like real distribution
While keeping content

GAN based SR Results

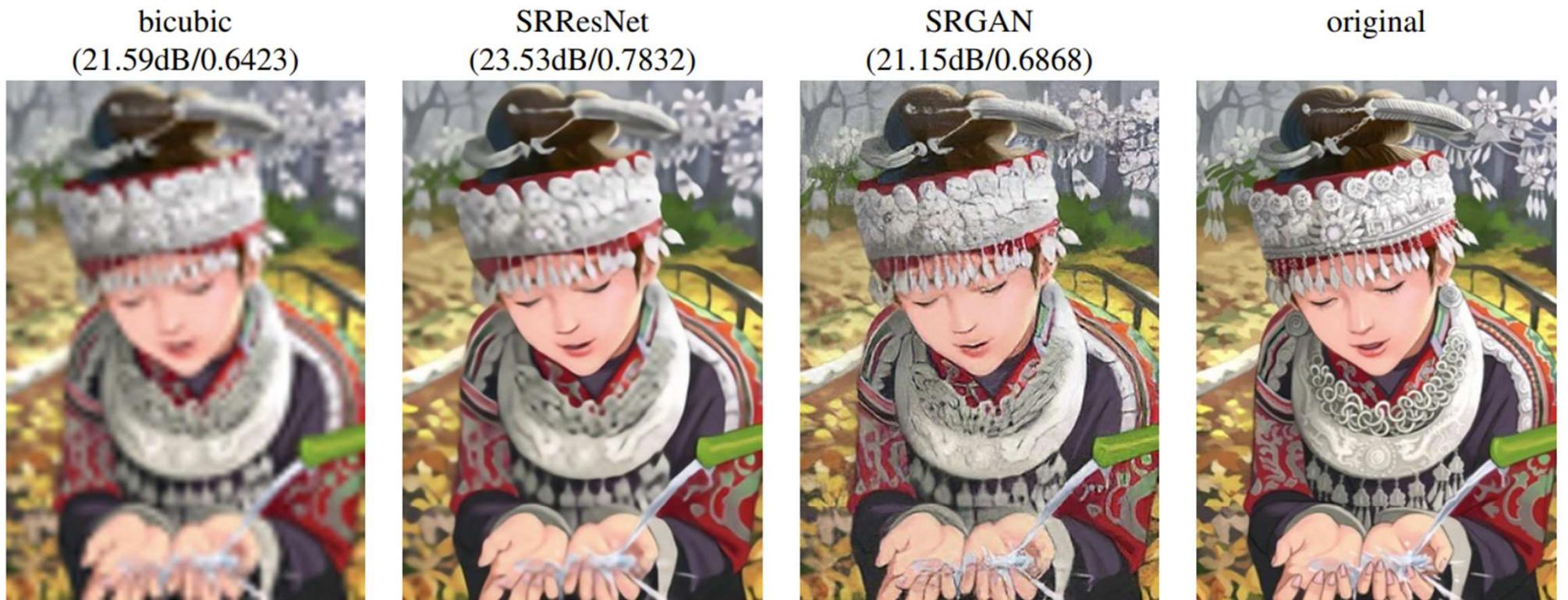
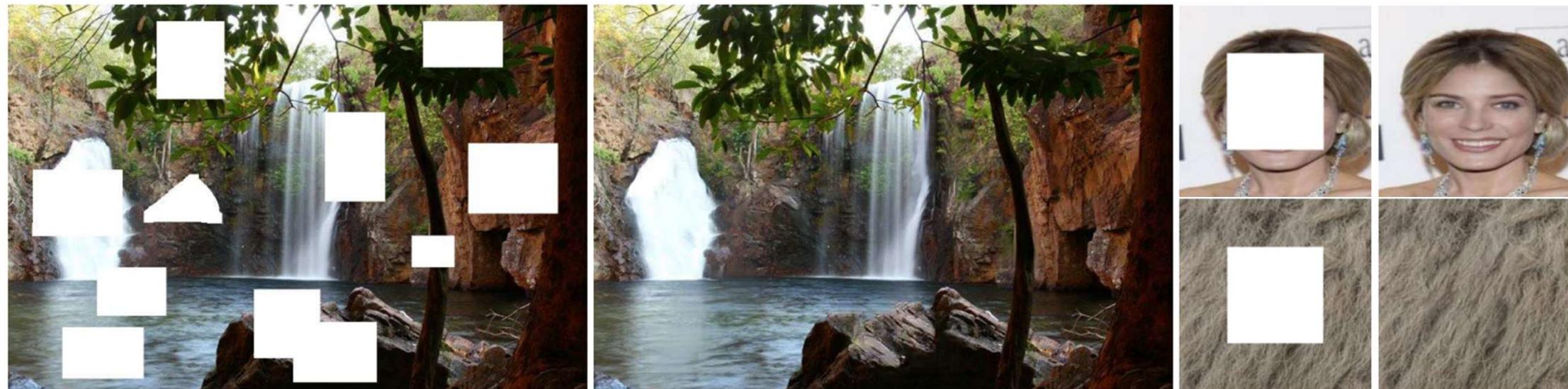


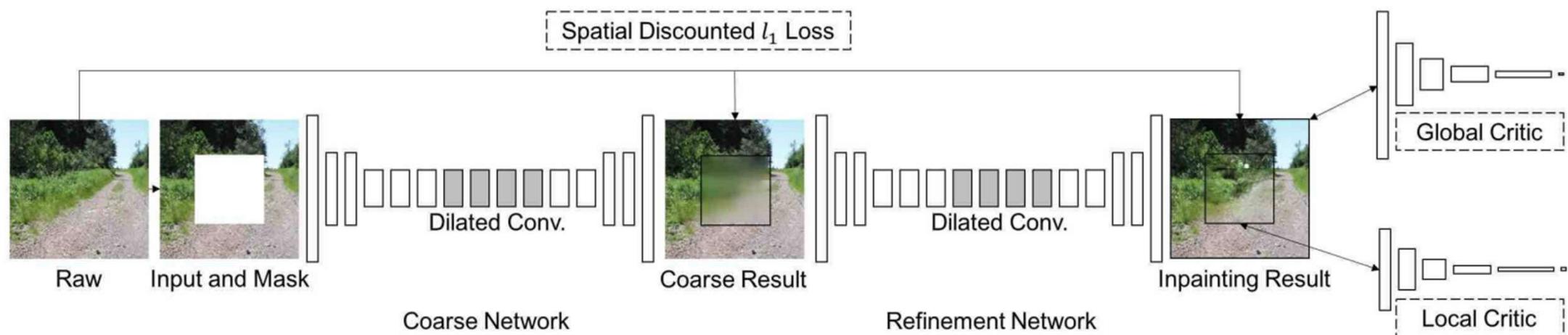
Figure 2: From left to right: bicubic interpolation, deep residual network optimized for MSE, deep residual generative adversarial network optimized for a loss more sensitive to human perception, original HR image. Corresponding PSNR and SSIM are shown in brackets. [4× upscaling]

The Methodology: Image Inpainting



Generative Image Inpainting with Contextual Attention, CVPR 2018.

The Methodology: Image Inpainting

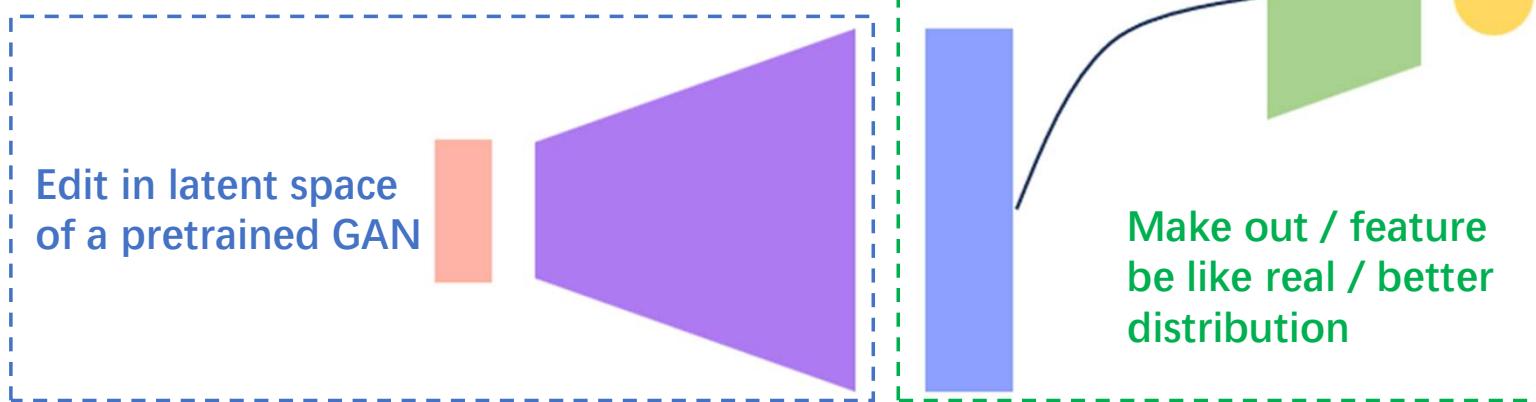


Make output be like real distribution
Global / Local Attention

Generative Image Inpainting with Contextual Attention, CVPR 2018.

Summary

Almost all GAN-based DIP
papers follow the
methodology (2015 — 2023)





中国科学技术大学

University of Science and Technology of China

谢谢观看！