

# 生成对抗网络 (GAN) 与遥感

Mia Feng

July 28, 2019

## 目录

- 判别模型 (D) 与生成模型 (G)
- 生成模型新秀: GAN
- GAN 与数据预处理: 数据融合、去噪
- GAN 与地球物理模型
- GAN 与模型优化: 假数据集扩充; 无监督数据集
- GAN 与模型泛化: 知识域迁移
- GAN 与预测: 从视频帧预测到台风云图预测

## 判别模型

对条件概率  $p(y|x)$  建模，意图在有限样本条件下，建立判别函数，实现直接预测。

例：已知羊的胡子长度 (beard\_length)、羊角大小 (horn\_size) 等特征，判别羊是山羊 (goat) 还是绵羊 (sheep)：

$$p(y_{\text{goat}}) = p(\text{goat} | x_{\text{beard\_length}}, x_{\text{horn\_size}}, \dots) \quad (1)$$

优点：

- 分类边界更灵活，适用于较多类别的识别；
- 判别模型的性能比生成模型要简单，较易学习

缺点：

- 不能反映训练数据本身的特性，变量间的关系不清楚，不可视。

针对  $y$  建模  $p(y)$  后，通过贝叶斯定律得到  $p(y|x)$  的概率。

例：已知山羊和绵羊的特征，如胡子长度 ( $\text{beard\_length}$ )、羊角大小 ( $\text{horn\_size}$ ) 等特征，判别羊是山羊 ( $\text{goat}$ ) 还是绵羊 ( $\text{sheep}$ ):

$$p(y_{\text{goat}}) = \frac{p(x_{\text{beard\_length}}, x_{\text{horn\_size}}, \dots | y_{\text{goat}}) p(y_{\text{goat}})}{p(x_{\text{beard\_length}}, x_{\text{horn\_size}}, \dots)} \quad (2)$$

优点:

- 模型的信息比判别模型丰富，单类问题中比判别模型灵活;
- 小数据集同样适用。

缺点:

- 学习和计算过程复杂。

## 关系

由生成模型可以得到判别模型，但由判别模型得不到生成模型。

## 生成模型的挑战

浅层生成模型：以高斯混合模型等为代表，采用浅层结构，由于其建模和表示能力有限，在遇到一些更复杂的涉及自然信号（比如人类语言、自然图像和视觉场景）的问题时就会遇到各种困难。

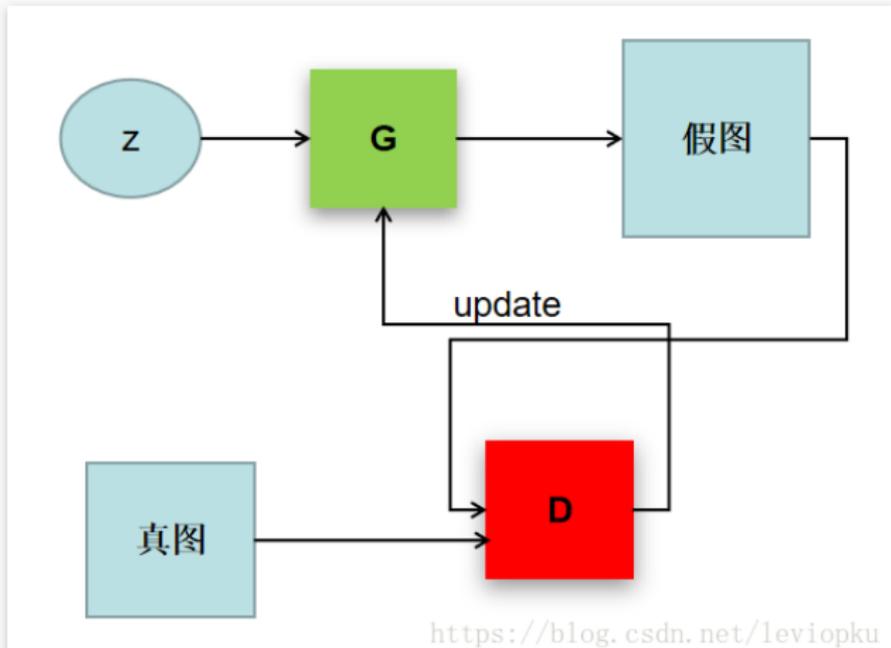
深度生成模型：对于处理似然函数中产生的复杂概率分布，以及平衡深度神经网络中的分段线性单元存在困难。

## 常用生成模型的特点

	Deep directed graphical models	Deep undirected graphical models	Generative autoencoders	Adversarial models
Training	Inference needed during training.	Inference needed during training. MCMC needed to approximate partition function gradient.	Enforced tradeoff between mixing and power of reconstruction generation	Synchronizing the discriminator with the generator. Helvetica.
Inference	Learned approximate inference	Variational inference	MCMC-based inference	Learned approximate inference
Sampling	No difficulties	Requires Markov chain	Requires Markov chain	No difficulties
Evaluating $p(x)$	Intractable, may be approximated with AIS	Intractable, may be approximated with AIS	Not explicitly represented, may be approximated with Parzen density estimation	Not explicitly represented, may be approximated with Parzen density estimation
Model design	Nearly all models incur extreme difficulty	Careful design needed to ensure multiple properties	Any differentiable function is theoretically permitted	Any differentiable function is theoretically permitted

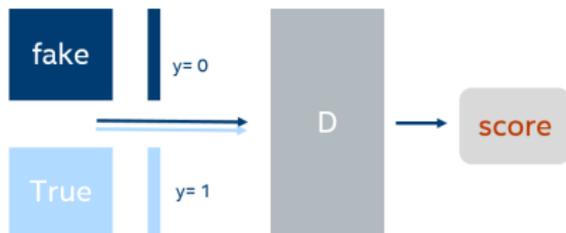
# GAN=G+D+repeat

- 生成器 G: generator, 负责凭空 (如随机生成一些噪声作为输入) 捏造数据。
- 判别器 D: discriminator, 负责判断输入的数据是不是真数据。
- 重复训练: 一般是先训练 D 再训练 G, 训练示意图如下:

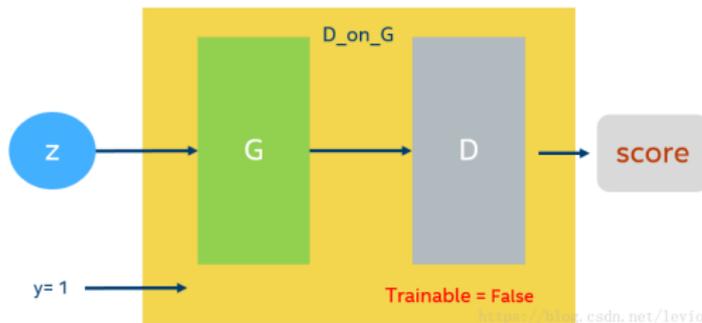


# 训练 GAN

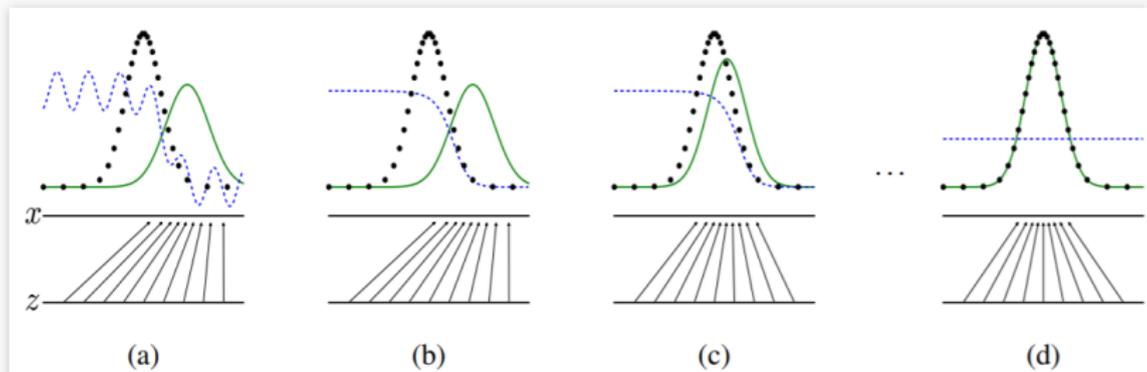
When training D:



When training G:



## GAN 训练实例



蓝色虚线为判别器，绿色线为生成器，黑色点线为真实数据生成器。  
从左到右为训练的演进中，分布的变化。最终 D 的损失控制在 0.5。

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

优点:

- 计算方面: 计算梯度只需 BP 而无需马尔科夫链; 模型可以融合多种函数;
- 统计方面: 生成器根据梯度流而非原始数据进行更新, 可以避免生成的数据严重相似于原始样本; 可以表示尖锐的或者退化的分布, 擅长处理复杂分布。

缺点:

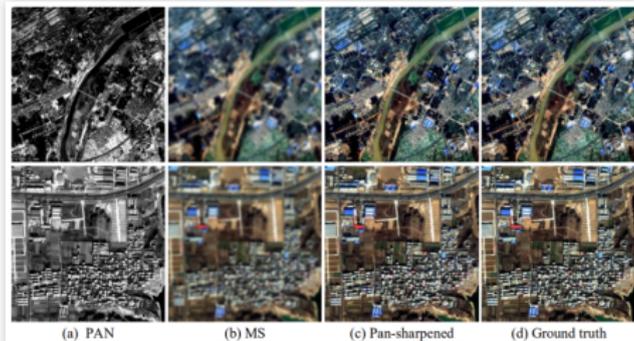
- 无显式的生成器所生成的数据的分布表达式;
- G, D 需要一起优化。

## GAN 与数据融合：谱分辨率、空间分辨率

Suarez et al, Deep Learning Based Single Image Dehazing, CVPARW2018.

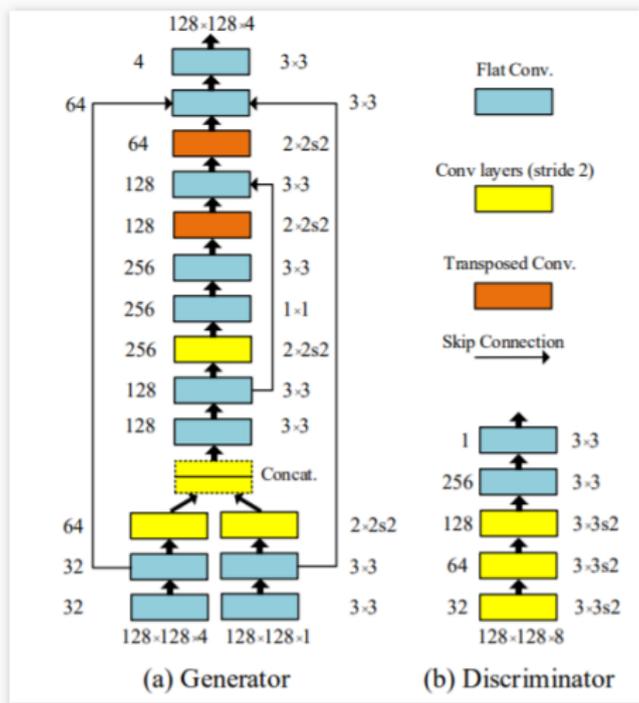
目的：将高空间分辨率、低光谱分辨率的影像与低空间分辨率、高光谱分辨率的影像融合得到高空间、光谱分辨率影像。

相比于原始的 GAN，主要不同在**网络结构**（模仿 UNet 的跳跃连接，采用双流结构）。损失函数与原始的 GAN 类似，但是不同于使用交叉熵度量，采用的 $\ell_1$  范数。



(a) 图为高空间分辨率影像，(b) 图为高光谱分辨率影像，(c) 图为融合后的影像，(d) 图为真实的高空间、高光谱分辨率影像。

# GAN 与数据融合：谱分辨率、空间分辨率



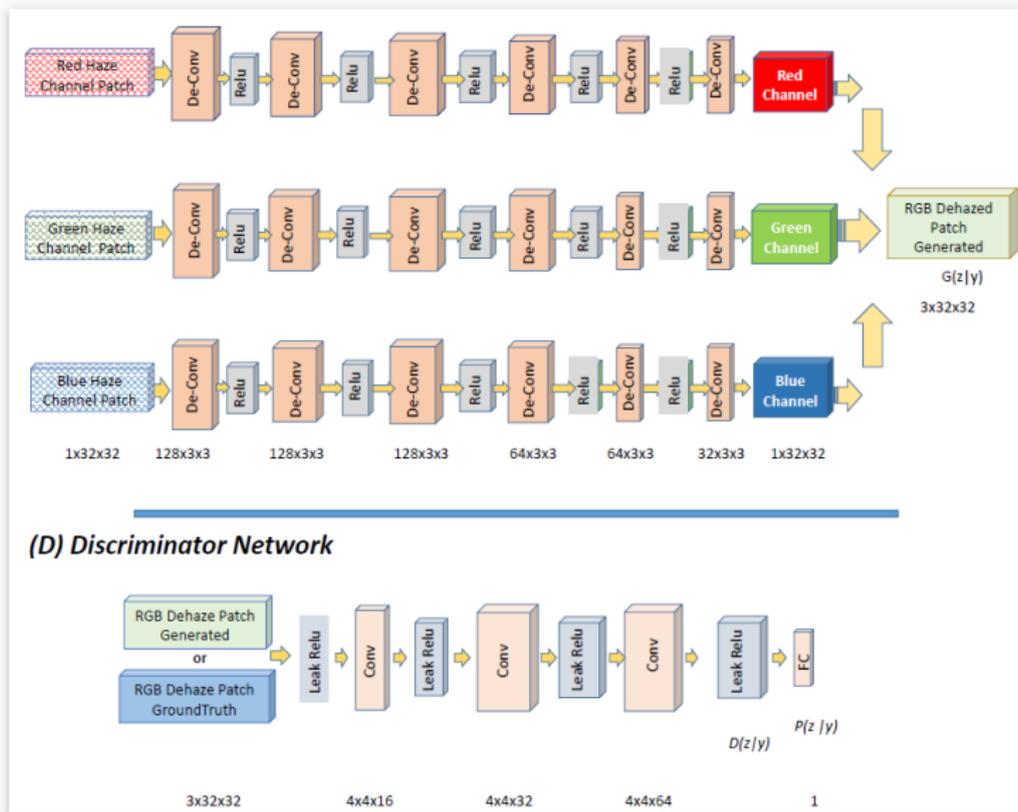
Liu et al, PSGAN, ICIP 2018.

目的：将高空间分辨率、低光谱分辨率的影像与低空间分辨率、高光谱分辨率的影像融合得到高空间、光谱分辨率影像。  
相比于原始的 GAN，主要不同在损失函数与网络结构（分波段进行）。

$$\mathcal{L} = 0.5\mathcal{L}_{\text{adversial}} + 0.2\mathcal{L}_{\text{Intensity}} + 0.2\mathcal{L}_{\text{SSIM}} + 0.2\mathcal{L}_{\text{quality}} \quad (3)$$

其中， $\mathcal{L}_{\text{adversial}}$  与原始 GAN 中的损失一致； $\mathcal{L}_{\text{Intensity}}$  衡量 RGB 色彩损失； $\mathcal{L}_{\text{SSIM}}$  衡量结构损失； $\mathcal{L}_{\text{quality}}$  衡量因为雾导致的图片质量损失。

# GAN 与影像去雾

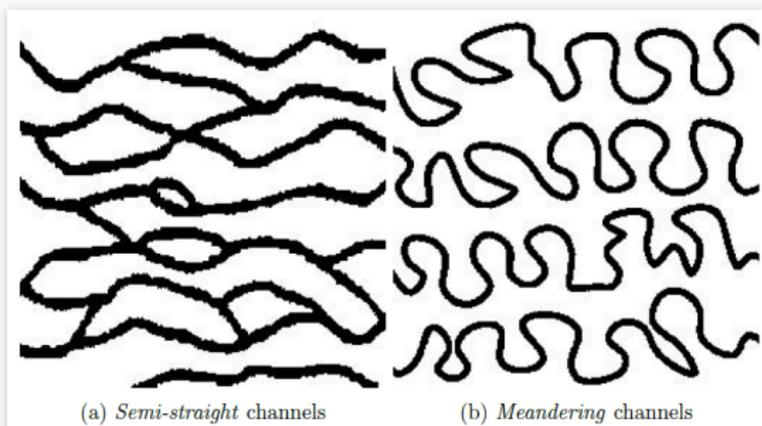


## GAN 辅助生成地球物理模型

Shing et al, Parametrization and generation of geological models with GAN, 2017

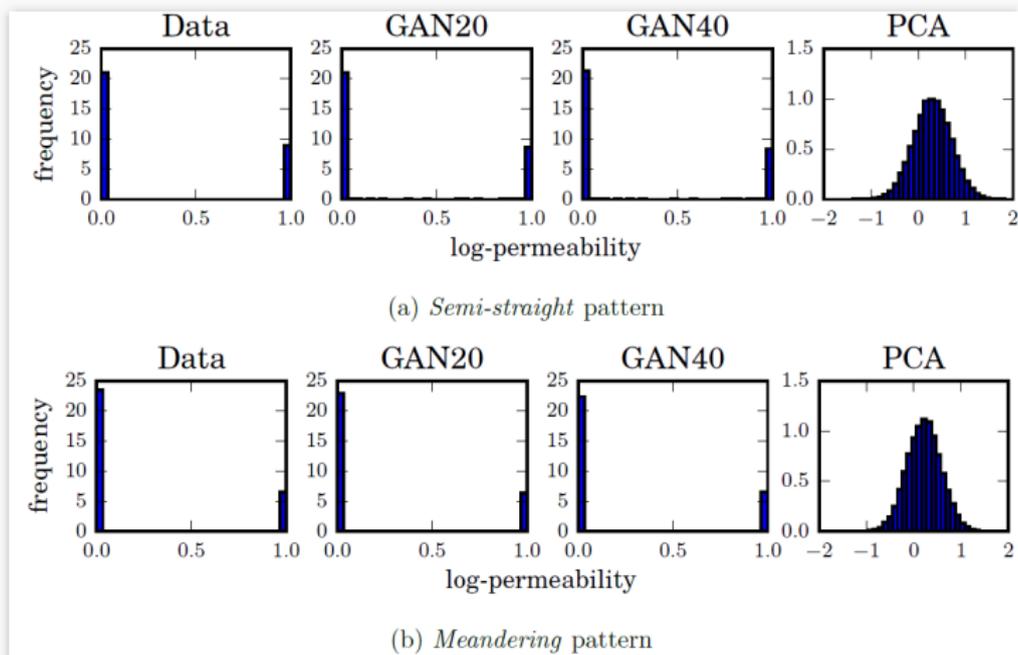
目标：使用 GAN 中的生成器模拟地球物理模型。

网络：Wasserstein GAN，与原始 GAN 相比，损失函数的衡量采用 EM 距离。EM 距离保证了产生数据的多样性，且可防止网络训练中的崩塌。



b 的生成复杂度高于 a

# 效果



图中为生成器生成的数据分布直方图，GAN20 与 GAN40 分别表示以 20 维噪声与 40 维噪声作为生成器的输入。

## 不确定性传播检验

假设一溢油问题遵循流体方程：

$$-\nabla \cdot (a\nabla p) = q \quad (4)$$

$$\varphi \frac{\partial s}{\partial t} + \nabla \cdot (sv) = q_w \quad (5)$$

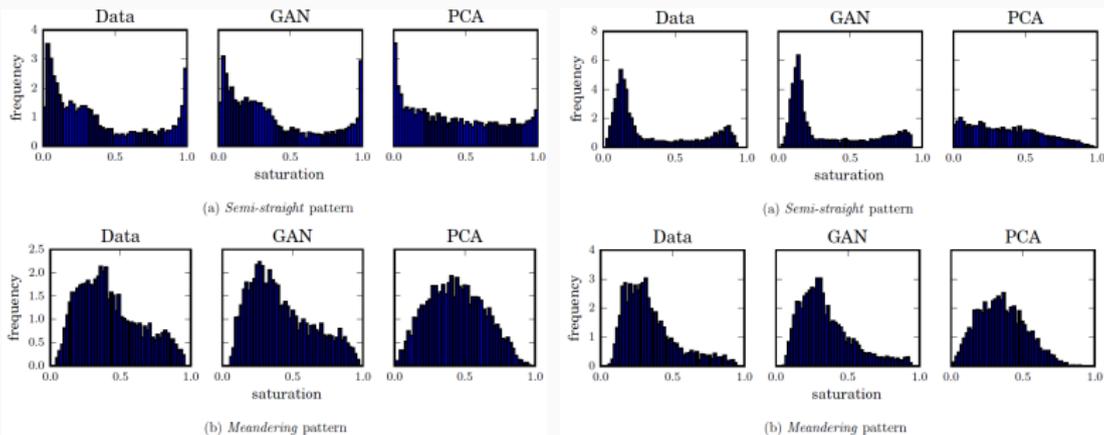
$p$  为流体压强， $q$  为流体源， $v$  为达西流速， $a$  为透过率， $s$  为水的饱和， $\varphi$  为多孔性。

*Quarter-five spot problem:* In this problem, injection and production points are located at  $(0, 0)$  and  $(1, 1)$  of the unit square, respectively. No-flow boundary conditions are imposed on all sides of the square. We assume unit injection/production rates, i.e.  $q(0, 0) = 1$  and  $q(1, 1) = -1$ .

*Uniform flow problem:* In this problem, uniformly distributed inflow and outflow conditions are imposed on the left and right sides of the unit square, respectively. No-flow boundary conditions are imposed on the remaining top and bottom sides. A total inflow/outflow rate of 1 is assumed. For the unit square, this means  $v \cdot \hat{n} = -1$  and  $v \cdot \hat{n} = 1$  on the left and right sides, respectively, where  $\hat{n}$  denotes the outward-pointing unit normal to the boundary.

不确定性研究基于上述两个问题进行。

# 不确定性传播效果 (GAN 指前文中的 GAN20)



(a) Quarter-five spot problem

(b) Uniform flow problem

Figure: 不确定性传播检验

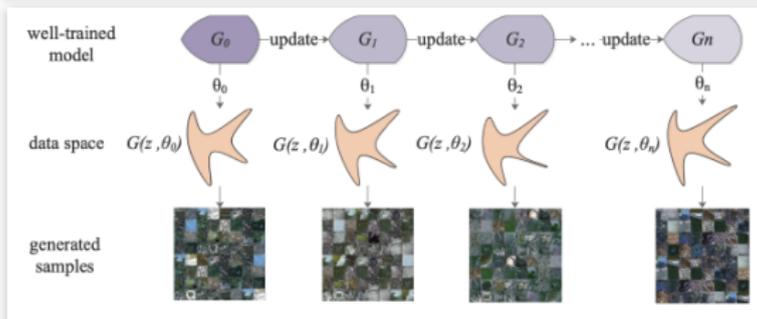
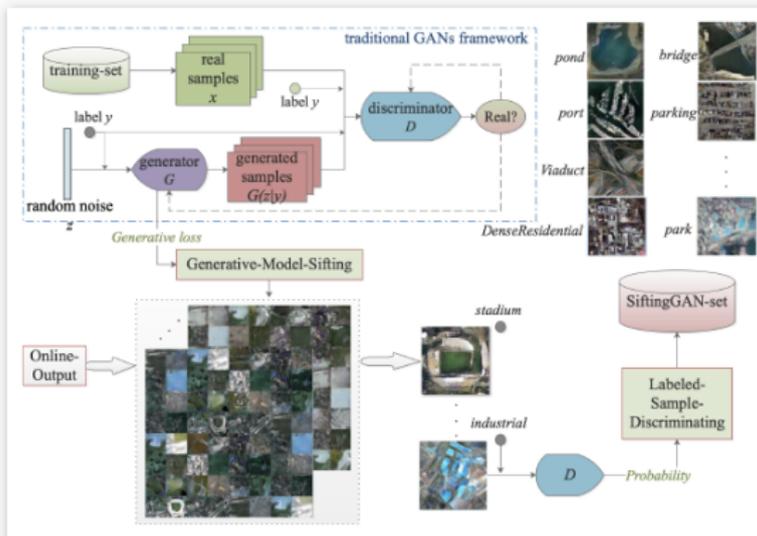
## GAN 生成扩充数据集

Ma et al, SiftingGAN, 2018

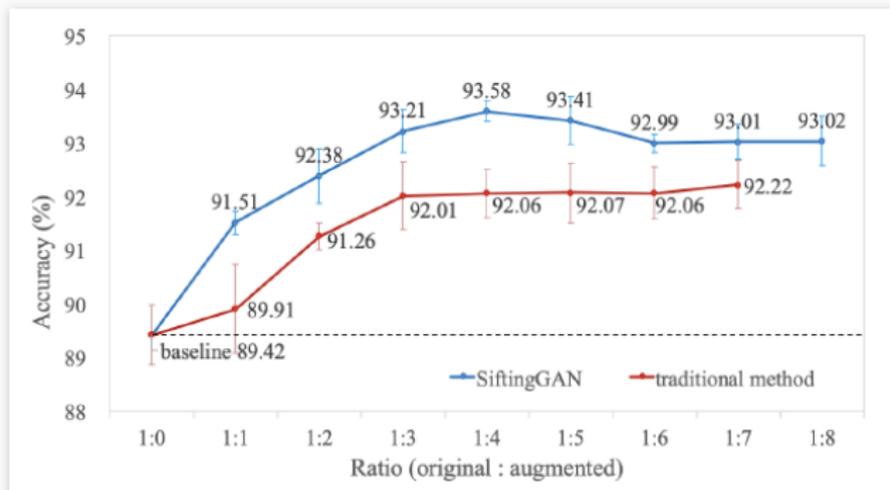
目标：验证 GAN 生成的假样本是否可以用于其他非 GAN 的网络结构，如 CNN，来提高分类准确率；GAN 模拟产生的扩充样本是否优于传统扩充手段（平移、翻转等）产生的样本。

区别：与原始 GAN 相比，设置了 G 和 D 的阈值，从而控制对生成的特定类别的假样本的接受率。这两个阈值分别产生了网络结构中所谓的 Generative model sifting 与 Labeled sample discriminating；同时为了保证 GAN 生成样本的速率，提出了在线 GAN 生成器（实际为在 G 优化中持续生成图片，而非对 G 优化完成后才生产图片）。

# GAN 生成扩充数据集



# 效果



相同扩充比率下，GAN 生成的样本所产生的的结果优于传统扩充方法所产生的结果。

Lin et al, MARTA GANs: Unsupervised Representation Learning for Remote Sensing Image Classification, 2017

目标：利用广泛存在的无标签数据集，提升分类结果。

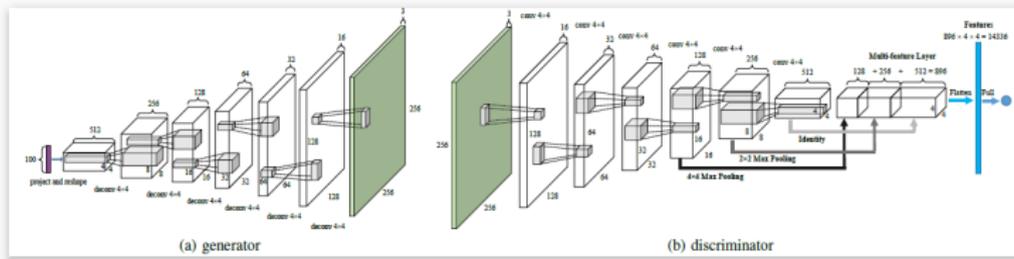
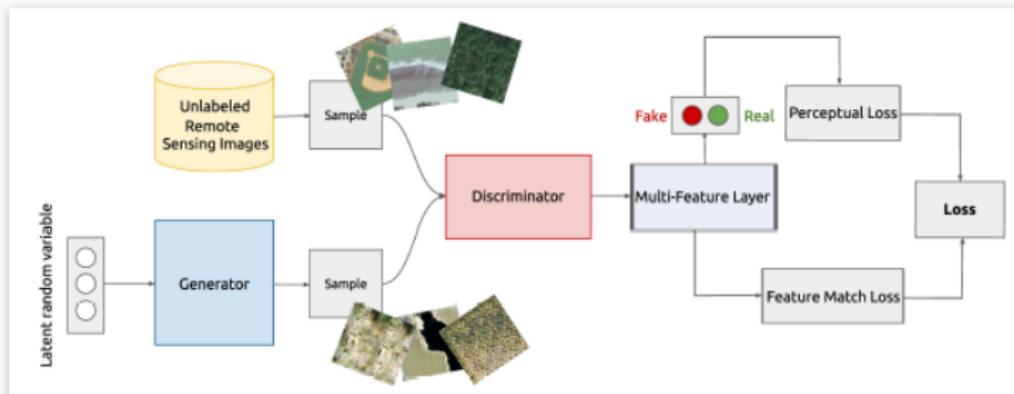
区别：与原始 GAN 相比，损失函数增加了  $\mathcal{L}_{\text{feature\_match}}$ ，用来衡量无标签样本与对应类别样本的一致性。结果上加无标签样本与生成器生成的假样本同时输入 D 中，得到的结果进行合并，输入到 Multi-Feature Layer 中。

$$\mathcal{L} = \mathcal{L}_{\text{adversarial}} + \mathcal{L}_{\text{feature\_match}} \quad (6)$$

$$\mathcal{L}_{\text{feature\_match}} = \|\mathbb{E}_{x \sim p_{\text{data}}(x)} [\log f(x)] - \mathbb{E}_{z \sim p_z(z)} [\log f(G(z))]\|_2^2 \quad (7)$$

f 为 Multi-feature Layer 中所表示的函数。

# GAN 利用无监督信息



# 效果

DataSet	Method	Description	Parameters	Accuracy
UC-Merced	SCMF [14]	Unsupervised	-	$91.03 \pm 0.48$
	UFL-SC [15]	Unsupervised	-	$90.26 \pm 1.51$
	OverFeat <sub>L</sub> + Caffe [4]	Supervised	205M	$99.43 \pm 0.27$
	GoogLeNet [5]	Supervised	5M	$99.47 \pm 0.50$
	<b>MARTA GANs</b>	<b>Unsupervised</b>	<b>2.8M</b>	$94.86 \pm 0.80$
Coffee	BIC [4]	Unsupervised	-	$87.03 \pm 1.07$
	OverFeat <sub>L</sub> + OverFeat <sub>S</sub> [4]	Supervised	289M	$83.04 \pm 2.00$
	CaffeNet [5]	Supervised	60M	$94.45 \pm 1.20$
	<b>MARTA GANs</b>	<b>Unsupervised</b>	<b>0.18M</b>	$89.86 \pm 0.98$

## GAN 与模型泛化：知识域迁移

Benjdira et al, Unsupervised Domain Adaptation Using Generative Adversarial Networks for Semantic Segmentation of Aerial Images, 2019

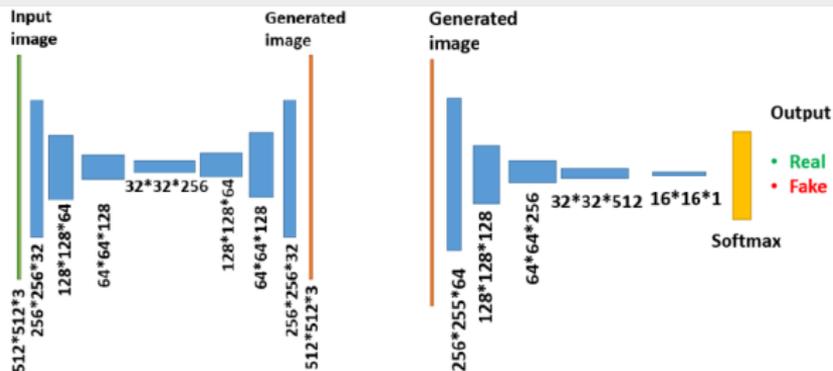
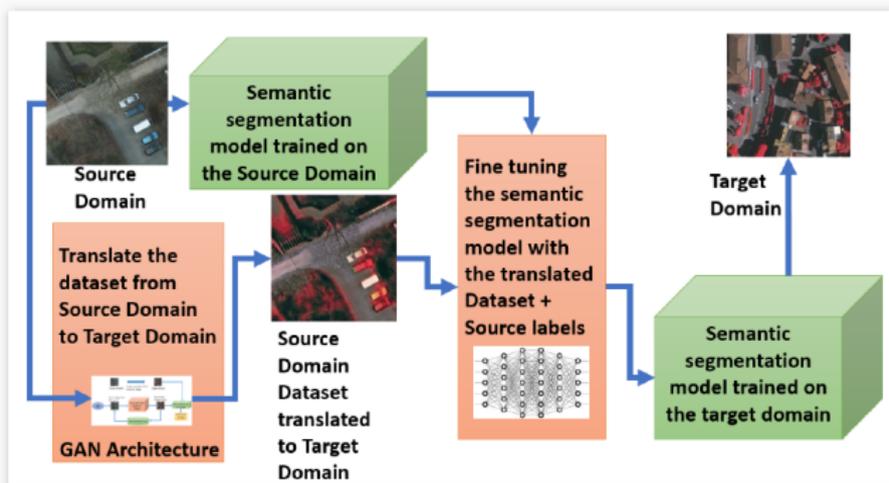
**目标：**源域与目标域的光谱波段、空间分辨率不同时，如何同时做到对源域与目标域的语义分割（如识别汽车所在的像素点）。

**区别：**与原始 GAN 相比，网络结构中 G 使用了自动编码-解码器；网络的损失中增加了重构损失  $\mathcal{L}_{\text{reconstruction}}$ ；网络流程上，将源域图像翻译到目标域图像后，再将其输入到之前在源域图像上调参好的语义分割模型进行调优，最终得到针对目标域的语义分割模型。

Domain shift factor	Efficiency	Examples of classes
Sensor	High	Building, Tree
Other factors	Conserve efficiency	Low vegetation
No Domain shift	Conserve efficiency	Cars

对于建筑物和树木的迁移分割结果较好；低矮植被和汽车的分割结果，

# GAN 与模型泛化：知识域迁移

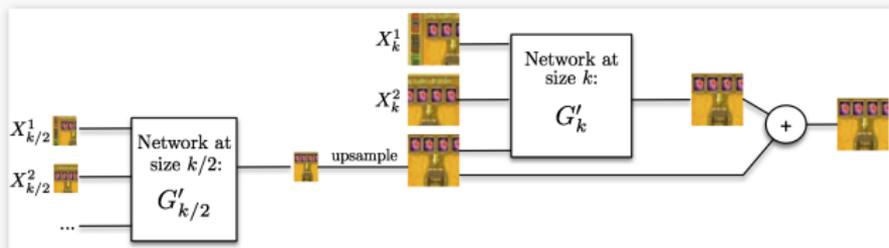


## GAN 与预测：从视频帧预测到台风云图预测

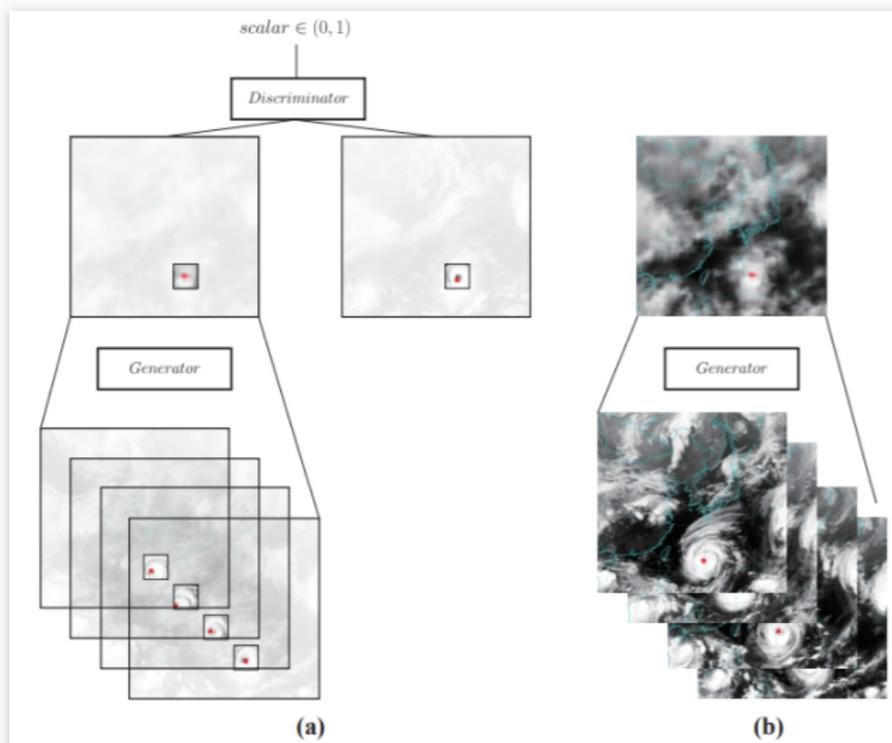
Mario et al, Typhoon track prediction using satellite images in a Generative Adversarial Network, 2018

目标：根据历史台风云图，预测未来时刻的台风云图（生成器的生成图片）。

区别：与原始 GAN 相比，网络结构中 G 使用了多尺度生成器，不同尺度的生成结果合并作为最终的生成器输出；网络的损失函数增加了梯度损失  $\mathcal{L}_{gdl}$  与特征匹配损失  $\mathcal{L}_{feature\_match}$ 。其网络模型来自于 Mathieu M, Couprie C, LeCun Y. Deep multi-scale video prediction beyond mean square error, 2015.



## GAN 与预测：从视频帧预测到台风云图预测



网络实际训练采用台风涡旋云系图片（ $32*32$ ），测试使用的完整云图图片（ $250*238$ ）。

# 效果

