

4.CNN & GAN

Convolutional Neural Network

- **灵魂拷问**：什么是全连接层？什么是卷积层？什么是池化层？(CNN,图像识别启动！！)

Answer: 全连接层 (Fully Connected Layer) 是将输入的所有神经元与输出层的所有神经元逐一连接，用于整合全局信息，常用于分类任务；卷积层 (Convolutional Layer) 通过卷积操作提取输入数据的局部特征 (如图像的边缘、纹理)，适合处理空间结构数据；池化层 (Pooling Layer) 则通过下采样 (如最大池化或平均池化) 减少数据维度，降低计算量并防止过拟合，同时保留重要特征。全连接层在图像处理中通常位于卷积层和池化层之后，它的作用是将前面层提取的特征综合起来，以执行分类、回归或其他任务。具体来说，全连接层可以将卷积层和池化层输出的特征图展平为一维向量，并通过权重矩阵和偏置向量进行线性变换，最终输出预测结果。

- 参数共享 (Parameter Sharing)

参数共享是卷积神经网络 (CNN) 中的一个重要概念，它指的是在卷积层中，同一个卷积核 (或滤波器) 的权重在不同的输入区域上被重复使用。这种机制显著减少了模型的参数数量，提高了计算效率，并使得模型能够捕捉到输入数据中的局部特征。

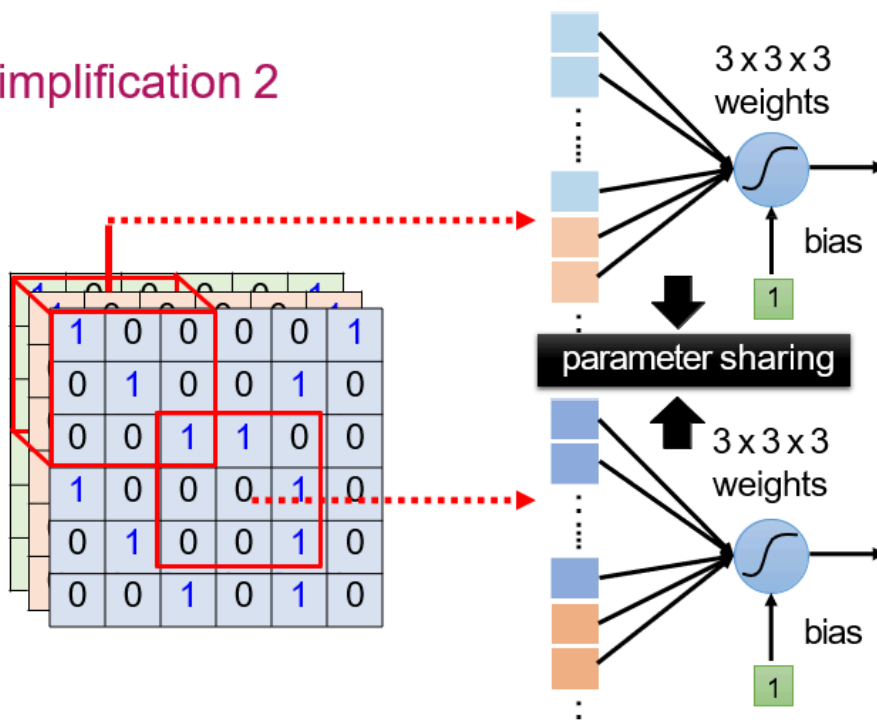
定义--->在卷积层中，每个卷积核的权重在整个输入数据上滑动 (或卷积)，并在每个位置上应用相同的权重。这意味着同一个卷积核的权重在不同的输入区域上是共享的。例如，如果一个卷积核的大小是 3×3 ，那么这个卷积核的 9 个权重在整个输入图像上被重复使用。

减少参数数量：通过参数共享，卷积层的参数数量显著减少。例如，对于一个 3×3 的卷积核，无论输入图像的大小如何，该卷积核只有 9 个权重。这使得模型更加紧凑，减少了计算和存储需求。

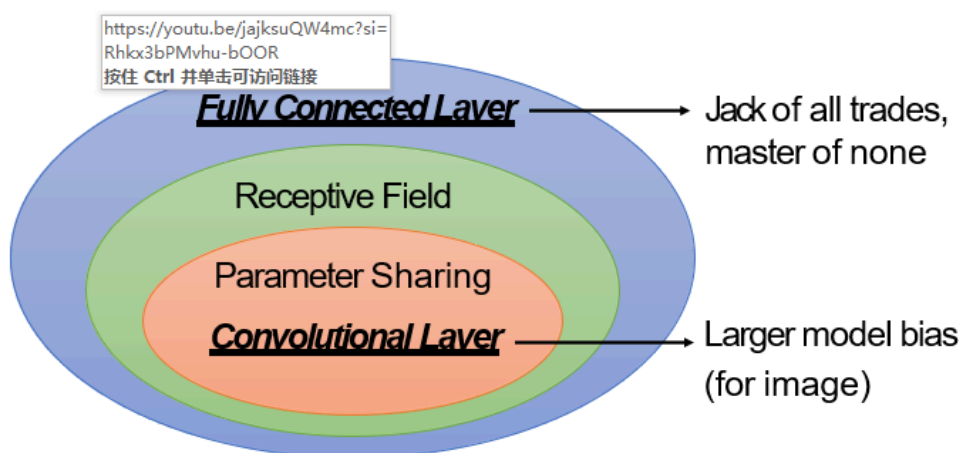
提高计算效率：参数共享使得卷积操作可以高效地在输入数据上滑动，减少了重复计算。这在处理大规模数据时尤为重要。

捕捉局部特征：参数共享使得卷积核能够捕捉到输入数据中的局部特征，如边缘、纹理等。这些局部特征对于图像识别等任务至关重要。

Simplification 2



Benefit of Convolutional Layer



- Some patterns are much smaller than the whole image.
- The same patterns appear in different regions.

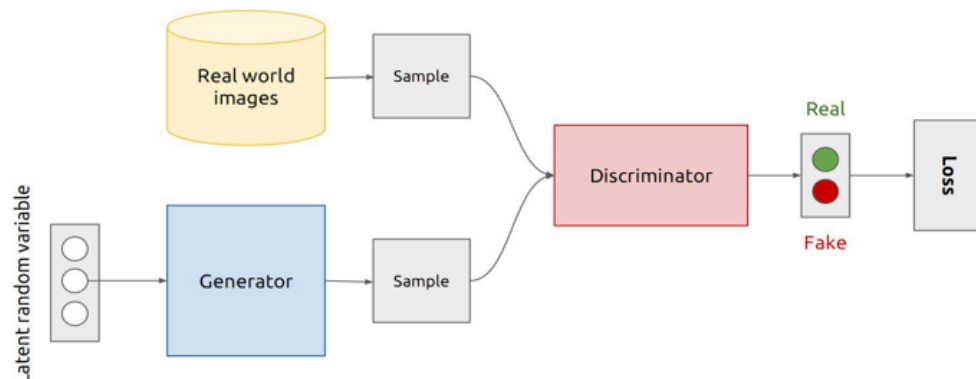
如上图所示，卷积层能够处理比整个图像小的多的模式。同时能够捕捉到出现在不同区域的相同模式。另外，卷积核（或滤波器）的值是模型的未知参数，这些参数在训练过程中通过反向传播进行学习和优化。这些参数的初始值通常是随机初始化的，然后在训练过程中逐渐调整，以捕捉输入数据中的局部特征

Generative Adversarial Networks

- 首先要知道这玩意儿广泛用于图像生成领域,也可以用于语句生成。属于无监督学习
- 生成对抗网络（GANs）是一种强大的生成模型，由 Ian Goodfellow 等人在 2014 年提出。GANs 通过两个神经网络——生成器（Generator）和判别器（Discriminator）——的对抗训练来生成逼真的图像。这种对抗训练过程可以被视为一个 minimax 两人游戏。

A minimax two-player game:

- Generator
- Discriminator



- 生成器（Generator）
生成器的目标是生成逼真的图像，这些图像在视觉上与真实图像难以区分。生成器通常从一个随机噪声向量 z 开始，通过一系列的神经网络层生成图像 $G(z)$ 。目标是生成的图像在视觉上与真实图像难以区分(learn the underlying data distribution)【而不是只为了最小化图片是假的概率】。
- 判别器（Discriminator）
判别器的目标是区分生成的图像和真实图像。判别器是一个二分类器，输出一个概率值，表示输入图像是真实图像的概率。输入：真实图像 x 或生成的图像 $G(z)$ 。输出：概率值 $D(x)$ ，表示输入图像是真实图像的概率。目标是最大化对真实图像的识别概率，同时最小化对生成图像的识别概率(Identify real data samples)。When the discriminator becomes unable to distinguish real from fake samples: The generator updates its weights.（当判别器无法区分生成的图像和真实图像时，更新参数。）
- 整个过程是什么样子的？
 - 1.初始化生成器和判别器：开始时，将 G 和 D 设置为神经网络。
 - 2.固定生成器 G ，更新判别器 D 。判别器被训练以区分真实图像（分配高分，例如 1）和生成的虚假图像（分配低分，例如 0）。
 - 3.固定判别器 D ，更新生成器 G 。生成器被更新，以生成能够“欺骗”判别器的图像。判别器的输出（例如 0.13）表示其置信度，生成器调整以使这一分数接近 1。
 - 4.过程2和3循环~~

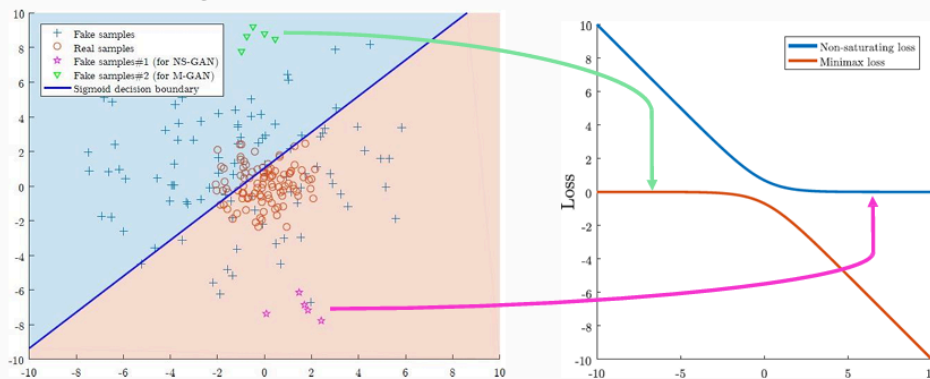
LSGAN

- 为什么要改进为LSGAN?
因为交叉熵损失存在问题如下:

Problem of Cross-Entropy Loss

Consequently: (when updating the generator)

- The non-saturating loss will cause almost no gradient for the fake samples in pink.
- The minimax loss will cause almost no gradient for the fake samples in green.

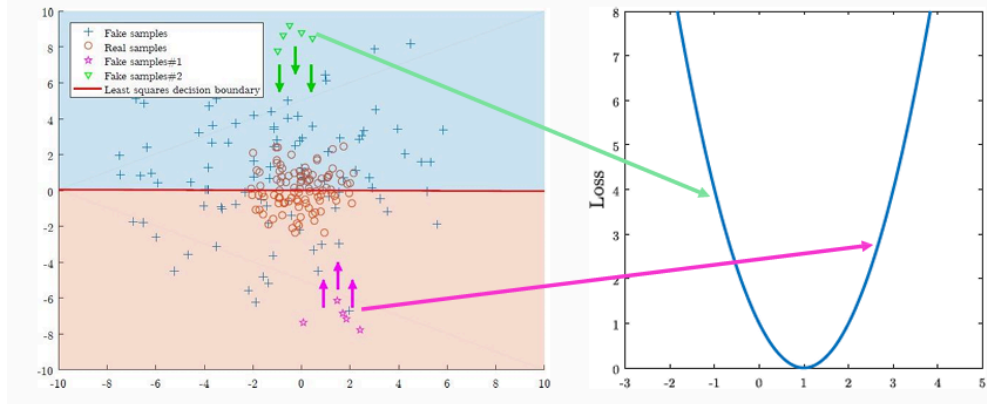


How to move the fake samples towards the real data?

- **非饱和损失 (Non-saturating Loss)** (粉色区域, 标记为假样本):
 - 当假样本 (粉色点) 位于粉色区域时, 非饱和损失会导致几乎没有梯度 (gradient) 供生成器更新。
 - 这意味着生成器无法有效学习如何改进假样本, 使其更接近真实数据。
- **最小最大损失 (Minimax Loss)** (绿色区域, 标记为假样本#2):
 - 当假样本 (绿色三角形) 位于绿色区域时, 最小最大损失同样会导致几乎没有梯度。
 - 生成器难以通过这种损失函数获得足够的反馈来调整假样本。
- 最小二乘法能有效改善。在一般分类任务中, 样本位置不变, 边界改变。在GAN中边界不变, 样本位置改变。(逆向过程?)

Intuition of Using Least Squares Loss

- The least squares loss is able to move the fake samples toward the decision boundary.

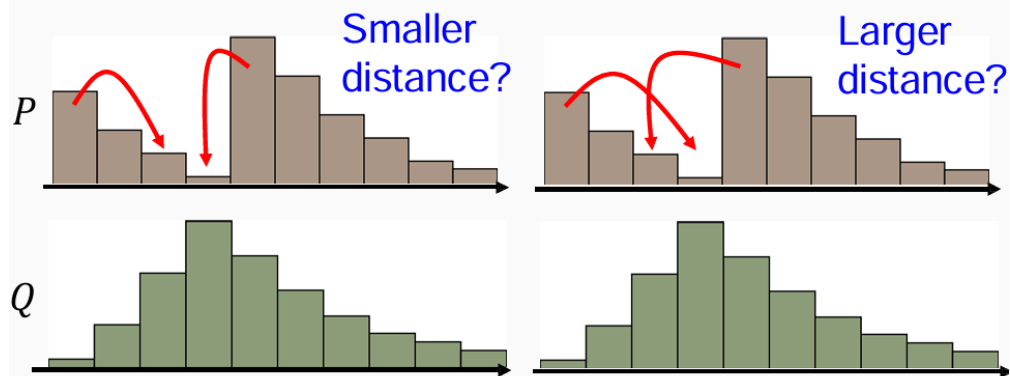


- LSGAN可以用于字符识别的数据增强再生成，也可以应用于场景图像去雨(Semi-MoreGAN)
 - LSGANs generate higher quality images than NS-GANs and WGANs-GP.
 - LSGANs perform more stably than NS-GANs.
 - LSGANs-GP is comparable with WGANs-GP in training stability.
 - LSGAN has many applications, including data augmentation and image deraining.

Evaluation of GANs

- Non-Saturating GAN Loss是关于最小最大损失(对于评分一直很低的情况会无法改进)的改进，更为稳定，input越大越饱和。对生成器的反馈越强。理解为它通过更稳定的权重更新机制避免了生成器饱和/参数更新停止。
- Frechet Inception Distance(FID)是一种评估生成质量的方式。使用 Inception 网络从中间层提取特征，它测量假图像分布和真实图像分布之间的距离，较低的 FID 值意味着更好的图像质量和多样性。
- Wasserstein distance
用于在Wasserstein GANs中测量真实样本和生成样本之间的相似性。
而Wasserstein GANs的设计目的是缓解GAN中的稳定性问题

Wasserstein distance



There are many possible "moving plans".

Using the "moving plan" with the smallest average distance to define the Wasserstein distance.

- 要注意GAN的训练并不稳定！在训练中存在Mode Collapse问题。即当生成器产生有限的样本变化。When the generator produces limited variations of samples

如何解决mode collapse?

- ① 梯度惩罚 (Gradient penalty)
- ② 更多样化的数据 (Collecting more varied training data)

如何改善GANs训练时的稳定性?

- ① Gradient penalty
- ② Batch normalization