

文章编号: 2095-4980(2023)02-0235-07

## 基于生成对抗网络的语义分割方法

刘可心

(电子科技大学 格拉斯哥学院, 四川 成都 611731)

**摘要:** 在语义分割模型结构不变的前提下, 为提升模型对图像分割的精确度, 引入生成对抗网络结构用于训练语义分割模型(SS-GAN)。SS-GAN 包含 3 个设计环节: 构建全卷积网络(FCN)结构的生成模型, 进行初步的图像分割; 设计具备像素间高阶关系学习能力的对抗模型, 提高生成模型的学习能力; 加入对抗损失辅助生成模型学习, 进一步促进生成网络自主学习像素间关系。在计算机视觉竞赛数据集(PASCAL VOC)和城市景观数据集(Cityscapes)上的实验结果表明, 引入生成对抗网络后取得了更好的效果, 2 个数据集的交并比(IoU)指标分别提高了 1.56%/1.17% 和 1.93%/1.55%。

**关键词:** 语义分割; 生成对抗网络; 全卷积神经网络; 空间连续性

**中图分类号:** TP391.4

**文献标志码:** A

**doi:** 10.11805/TKYDA2020554

## A method by Generative Adversarial Network in semantic segmentation

LIU Kexin

(School of Glasgow, University of Electronic Science and Technology of China, Chengdu Sichuan 611731, China)

**Abstract:** In order to improve the accuracy of image segmentation without changing the structure of original semantic segmentation models, an approach is proposed to train Semantic Segmentation models by using Generative Adversarial Network(SS-GAN). There are three steps related to this work: constructing the generative model of Fully Convolutional Network(FCN) structure to segment image preliminarily; constructing the adversarial model which can learn the high-order relationship between pixels and training it to improve the learning ability of generative model; adding the anti-loss to assist generative model training, encouraging generative network to learn the relationship between pixels independently. Experiments on Pattern Analysis, Statistical Modeling and Computational Learning (PASCAL VOC) and Cityscapes datasets show that the proposed method achieves better performance than the existing advanced methods, and improves Intersection over Union(IoU) by 1.56%/1.17% and 1.93%/1.55%, respectively.

**Keywords:** semantic segmentation; Generative Adversarial Network; Fully Convolutional Network; spatial contiguity

目前最先进的语义分割技术均利用卷积神经网络实现。2014 年, Long 等<sup>[1]</sup>通过对 AlexNet<sup>[2]</sup>、GoogleNet<sup>[3]</sup>、视觉几何组网络(Visual Geometry Group, VGG)<sup>[4]</sup>等高精度的分类神经网络进行结构改造, 提出了全卷积网络(FCN), 实现了端到端的图像语义分割, 并在计算机视觉竞赛数据集(PASCAL VOC)、NYUDv2<sup>[5]</sup>、SIFT Flow 等多个数据集上精确度得到较大提升。此后, 基于 FCN 结构设计的语义分割网络大量涌现: 2016 年, Chen 等提出了 DeepLab 系列<sup>[6]</sup>, 通过在网络中引入条件马尔可夫随机场(Conditional Random Field, CRF)、空洞卷积等方法, 进一步提高了语义分割的精确度; 2015 年, Ronneberger 等提出 U-Net<sup>[7]</sup>, 在 FCN 结构的基础上设计了编码器-解码器结构完成语义分割。然而, 无论是 FCN 结构的分割模型还是 U-Net 分割模型, 在模型训练中预测像素标签时, 均未考虑像素间的相关性, 这种相关性常用各种后处理技术来弥补。但后处理技术实现困难复杂, 运行速度较慢, 很难运用在视频分割或实时分割中。

收稿日期: 2020-10-22; 修回日期: 2020-12-23

为在去除后处理技术的同时保持分割精确度，本文利用 Goodfellow 等<sup>[8]</sup>提出的生成对抗网络(GAN)挖掘像素间的高阶联系。通过加入对抗网络，在不改变原有分割网络结构、参数量、计算量的前提下，辅助生成模型学习像素间的高阶关系。通过在 Cityscapes 数据集和 PASCAL VOC 数据集的实验证明，生成对抗网络能够辅助模型有效提升分割精确度。

## 1 SS-GAN 模型

### 1.1 SS-GAN 结构设计

整个结构逻辑如图 1 所示。将训练图片送入 Generative network(下称 G)，G 提取输入图中有效的语义信息和特征，并生成对应的分割结果；然后将标注的分割结果和 G 生成的分割结果分别送入 Adversarial network(下称 A)，A 包括 extractor 和 convnet。extractor 进一步对分割结果进行特征提取，形成更高阶的特征向量。同时将原图再送至 convnet 中进行粗提取，以补充原图在经由 G 和 A 的 extractor 两次提取后损失的原始信息。将分割结果生成的高阶特征向量和 convnet 生成的特征向量进行连接，并用全连接层加以整合。最后经由 SigMoid 层判定当前输入至 A 的分割结果是属于 G 生成的分割结果，还是原始标注的分割结果。根据 A 的输出结果对 G 的权重进行再调整，最后使 A 完全无法区分当前输入 A 的分割结果是 G 生成的结果还是原始标注，以提高 G 生成的分割效果。

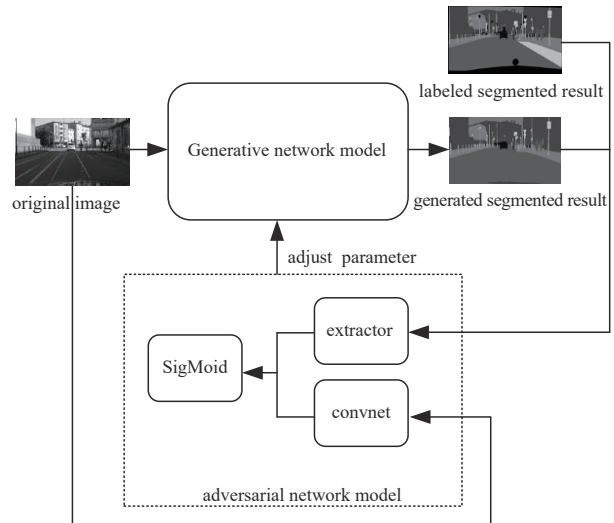


Fig.1 SS-GAN network structure  
图 1 SS-GAN 网络结构

### 1.2 生成网络模型

SS-GAN 生成模型采用 VGG16 和残差网络 ResNet50 作为基础模型。VGG 网络(图 2)的结构设计简洁规整，易于调整，在分类和检测任务中都取得了优秀的性能，在 2014 年的 ImageNet Large Scale Visual Recognition Challenge(ILSVRC)比赛中，VGG 在 Top-5 中取得了 92.3% 的正确率。而 ResNet 网络(图 3)通过残差结构这一精巧设计，解决了网络层级较深时容易出现的梯度爆炸和梯度消失问题，其网络设计简单，容易收敛，普遍适用于各种视觉问题，成为近年来各种比赛中基础模型之一。VGG16-G 与 ResNet50-G 的详细网络结构对比如表 1 所示，其中 G 表示在分割模型中引入了生成对抗网络结构。

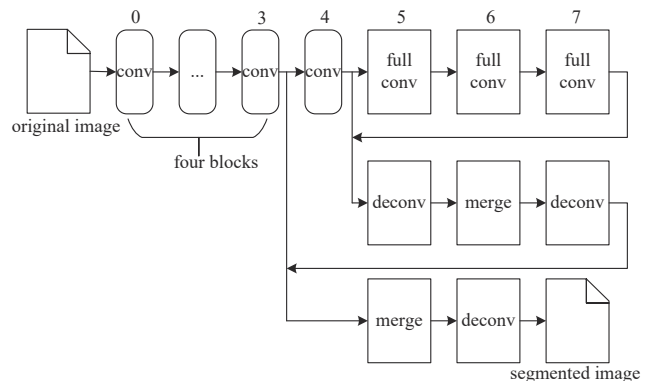


Fig.2 VGG16 network structure  
图 2 VGG16 结构

传统 VGG16 模型为图像分类模型，在模块 4 之后均为 3 层全连接层。而特征图在进入全连接层之前，会进行一次维度变换，将 y 轴的特征向量连接在一起，从特征图变为空间向量。这样有利于语义特征的提取，但破坏了对象的空间结构特征，而语义分割需要同时获取对象的语义特征和空间特征。为保持特征图的空间特征，将网络的全连接层转变成全卷积层，全连接层的权重通过变形转换为对应全卷积层的卷积核参数。如对于模块 5 的全卷积层，对应先前全连接层的权重形状为(25 088,4 096)。由于全连接层输入的特征向量形状为(25 088)，由形状为(7,7,512)的特征图变形而来，可以设置全卷积层的卷积核尺寸为(7, 7)，输入通道为 512，输出通道为 4 096，再将全连接层的权重进行形状变换后嫁接过来。第一层反卷积所接收的特征图相较于原图放大了 32 倍，如果直接使用该特征图进行反卷积放大原图，分割效果较差。因此通过反卷积放大后与第 4 个模块输出的特征图进行融合构成新的特征图。以此类推，将与第 3 个模块输出的特征图融合后的新特征图进行 8 倍插值放大，获得更精细的分割效果。

ResNet50 在不增加网络参数量的前提下，将第 3 和第 4 两个模块中的标准卷积核替换为空洞卷积核，以增加特征图的感受野范围。感受野扩大使深层特征图中的每一个像素均可获取浅层特征图中更多的信息，分割的精细程度得到提高。同时引入全局平均池化层增强空间金字塔池化结构(Atrous Spatial Pyramid Pooling, ASPP)。该

层为特征全图的全局平均池化层，对一张特征图上所有像素点取平均值，使原特征图成为特征向量，再由卷积层特征重组后重新放大，与其他特征图融合。

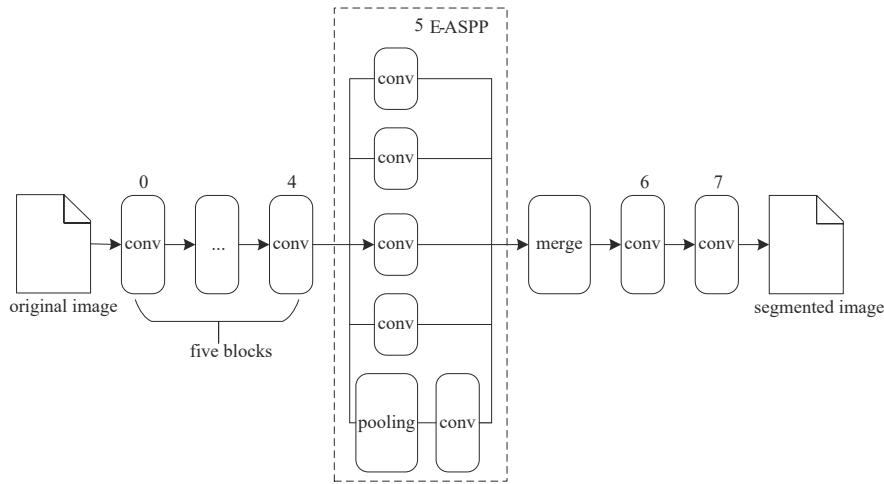


Fig.3 ResNet50 network structure  
图3 ResNet50 结构

表 1 生成模型结构

Table1 Generated model structure

NO.	VGG16-G		ResNet50-G	
	type of layer	parameter of layer	type of layer	parameter of layer
0	convolution/pooling	$[3 \times 3, 64] \times 2$	convolution/pooling	$[7 \times 7, 64] \times 1$
1	convolution/pooling	$[3 \times 3, 128] \times 2$	convolution	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
2	convolution/pooling	$[3 \times 3, 256] \times 3$	convolution	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
3	convolution/pooling	$[3 \times 3, 512] \times 3$	hole convolution	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
4	convolution/pooling	$[3 \times 3, 512] \times 3$	hole convolution	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
5	convolution	$[7 \times 7, 4096] \times 1$	enhance ASPP	-
6	convolution	$[1 \times 1, 4096] \times 1$	convolution	$[1 \times 1, 256] \times 1$
7	convolution	$[1 \times 1, classes] \times 1$	convolution	$[1 \times 1, classes] \times 1$

### 1.3 对抗网络模型

在对抗模型中，为保证对抗模型对像素间高阶关系的学习，与一般对抗模型只输入生成结果与真实结果不同，原图同时也被输入到模型当中，通过特征提取器进行特征提取后与分割图特征融合(图4)。

由于对抗模型需要判断分割图的来源，而分割图本身空间结构简单，语义信息高度概括，若直接输入真假分割图进行对抗训练，容易导致对抗模型通过细节区别快速收敛，而未学到像素间高阶关系。将输入的真实分割图进行 one-hot 变换成一个  $H \times W \times classes$  的特征图，对于每个像素，描述它的特征向量仅在像素所属类别上维度值为 1，其余为 0，而生成模型传递的生成分割图也为一个  $H \times W \times classes$  的特征图，该特征图在  $classes$  维度上所有数值均为 0~1 之间的浮点数。对抗模型仅凭判断输入分割特征图中是否含有 0 和 1 即可判断当前特征图的来源。为此，一方面增加对于原始图片特征的提取，利用特征提取器(如 VGG/ResNet 中的前几层)对图片进行特征提取，得到图片的一类低级特征。同时由于分割图是经过提炼较为抽象的特征图，使用深层网络会使得信息损失更加严重，因此使用一个较浅的小网络对其进行特征变换，得到第二类低级特征。然后将两类低级特征进行合并，再使用卷积神经网络进行特征融合。另一方面，通过对真实分割图的 one-hot 特征图进行放缩变化，将 0、1 的特征向量转变为浮点数的特征向量。对于真实分割图中的像素  $i$ ，假设其 one-hot 的特征向量为  $v_i$ ，进行放缩

后的特征向量为  $\bar{\mathbf{v}}_i$ 。则设定一个固定值  $\varepsilon$ ，表示特征向量中值为 1 的向量维(假设位置为  $l$ )在进行放缩变化后值最低不能小于  $\varepsilon$ 。同时获得生成分割图中相同位置像素的特征向量  $\mathbf{u}_i$ ，定义  $\bar{\mathbf{v}}_i$  中位置为  $l$  的向量值为：

$$\bar{\mathbf{v}}_i^l = \max(\varepsilon, \mathbf{u}_i^l) \quad (1)$$

而  $\bar{\mathbf{v}}_i$  中其他位置  $c$  的向量值为：

$$\bar{\mathbf{v}}_i^c = \frac{\mathbf{u}_i^c(1 - \bar{\mathbf{v}}_i^l)}{1 - \mathbf{u}_i^l} \quad (2)$$

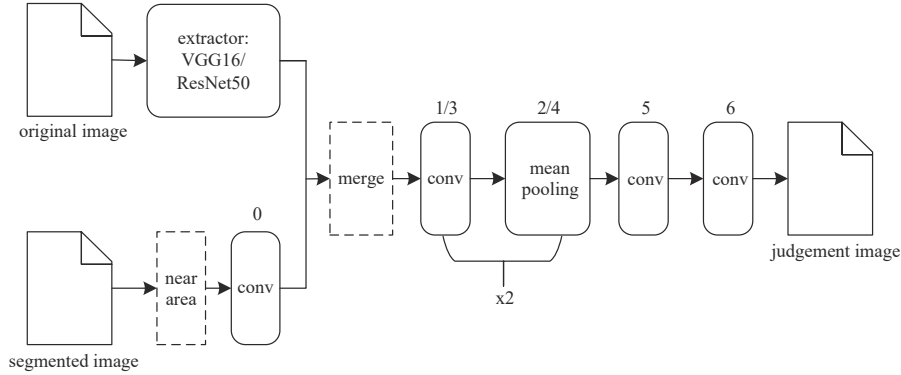


Fig.4 Structure of the adversarial model

图 4 对抗网络模型结构

通过以上的放缩变换，当生成分割图中的像素类别预测正确，且概率超过  $\varepsilon$  时，真实分割图中相同位置的像素所变换的特征向量将与生成分割图中完全相同；若低于  $\varepsilon$ ，则按照比例为其他维度增加概率值。通过这两方面操作，进一步缩小真实分割图与生成分割图间的显著区别，增大对抗网络区分真假样本的难度，帮助生成模型寻找像素间的空间连续性。此外，为进一步增强对抗模型的鲁棒性，在最后一层将特征图的尺寸转化为  $4 \times 4$  的小格评判图，每一个小格代表当前区域属于真实分割图还是合成分割图的概率。对每一个小格单独进行归一化操作，以防止当分割图仅在某一部分出现严重偏差时对特征图全局进行梯度调整。对抗模型结构如表 2 所示。

表 2 对抗模型结构表

Table 2 Structure of the adversarial model

NO.	type of layer	parameter of layer
0	convolution	$[3 \times 3, 64] \times 1$
1	convolution	$[3 \times 3, 128] \times 1$
2	average pooling	$2 \times 2, stride=2$
3	convolution	$[3 \times 3, 256] \times 1$
4	average pooling	$2 \times 2, stride=2$
5	convolution	$[3 \times 3, 512] \times 1$
6	convolution	$[3 \times 3, 1] \times 1$

#### 1.4 损失函数

在训练中需要对生成网络和对抗网络分别进行训练，以达到互相对抗、互相调整的目的。交叉熵损失函数对误差敏感，作用于反向传播能使模型快速收敛。基于交叉熵损失函数，分别设计用于生成模型训练的损失函数  $loss_g$  和对抗模型训练的损失函数  $loss_d$ 。当训练某一方模型时，需将另一方模型的权重固定，以保证利用模型之间相互竞争交替对抗增长。

生成模型损失函数为：

$$loss_g = loss_{mcc}(y, g(x)) - loss_{bcc}[0, d(g(x), x)] \quad (3)$$

式中： $loss_{mcc}(y, g(x))$  为单独训练生成模型损失函数； $loss_{bcc}(0, d(g(x), x))$  为引入对抗模型辅助训练的生成对抗损失函数。

$loss_{mcc}(y, g(x))$  为多类别交叉熵损失函数，其中  $y$  表示原图  $x$  中像素值所属类别的 one-hot 形式， $g(x)$  为生成模型对应的分割结果。 $loss_{mcc}(y, g(x))$  定义为：

$$loss_{mcc}(y, g(x)) = -y \ln g(x) \quad (4)$$

在将生成模型训练至收敛后，引入对抗模型对生成模型进一步辅助训练。

$-loss_{bcc}[0, d(g(x), x)]$  为基于二值交叉熵损失函数推演出的生成对抗损失项，其中  $d(g(x), x)$  为对抗模型输出的结果。训练生成模型时，首先将图像  $x$  输入到生成模型中，生成的分割结果  $g(x)$  与图像  $x$  一同输入到对抗模型中，根据对抗结果  $d(g(x), x)$ ，使用二值交叉熵损失函数计算得到生成对抗损失项。其中，二值交叉熵损失函数定义为：

$$loss_{bce}[x, d(g(x), x)] = -\{x \ln d(g(x), x) + (1-x) \ln [1-d(g(x), x)]\} \tag{5}$$

由于生成模型的任务是欺骗当前对抗模型，使其无法区分输入对抗模型中的分割图片来源，从而对  $g(x)$  的标签为 1(表示来源为数据集)，因此对于对抗模型的损失函数项为  $+loss_{bce}[1, d(g(x), x)]$ ，而根据 Goodfellow<sup>[8]</sup>中给出的结论，对抗模型在判断分割图片的来源是真实还是合成时，使用  $-loss_{bce}[0, d(g(x), x)]$ 能让生成模型的梯度下降更为稳定，因此采用其作为生成对抗损失项。

当训练对抗模型时，对于由生成模型合成的分割图片，其标签为 0，而从数据集中真实取得的分割图片，其标签为 1。期望对抗网络在输入的分割图片为真实的情况下，预测值能够接近 1，而在输入的分割图片为合成的情况下，预测值能够接近 0。故训练对抗模型的损失函数定义为：

$$loss_d = loss_{bce}[1, d(y, x)] + loss_{bce}[0, d(g(x), x)] \tag{6}$$

对于训练集中含有  $N$  张图片，第  $x_i$  张图片对应的分割图片  $y_i$ ，整个模型的损失函数定义为：

$$loss = \frac{1}{N} \sum_{i=1}^N \lambda loss_d \tag{7}$$

式中  $\lambda$  为超参数，用于调节对抗网络在训练初期时的损失大小。初期生成模型生成效果已经较好，而对抗模型由于仍处于初始化状态，提供的损失较大，需要通过超参数对其进行调整，以减小对抗损失对于生成模型调整的梯度大小。

## 2 实验

### 2.1 语义分割数据集

PASCAL VOC 2012 数据集中共有 12 031 张图片，拥有语义分割标签的训练图片 1 456 张，测试图片 1 449 张，所含语义分割类别为 20 类，同时使用拥有 10 582 张的 SBD 语义分割数据集进行数据集补充。Cityscapes 为包含来自 50 个不同城市的街道场景中记录的多种立体视频序列的大型数据集，拥有 5 000 个高质量像素级标注图片和 20 000 个粗糙标注图片。实验中，训练集图片数量为 3 475 张，测试集图片数量为 1 525 张，图片中物体类别共有 30 类，语义分割任务中选取其中 19 类。PASCAL VOC 2012 数据集与 Cityscapes 数据集的部分数据样例见图 5。

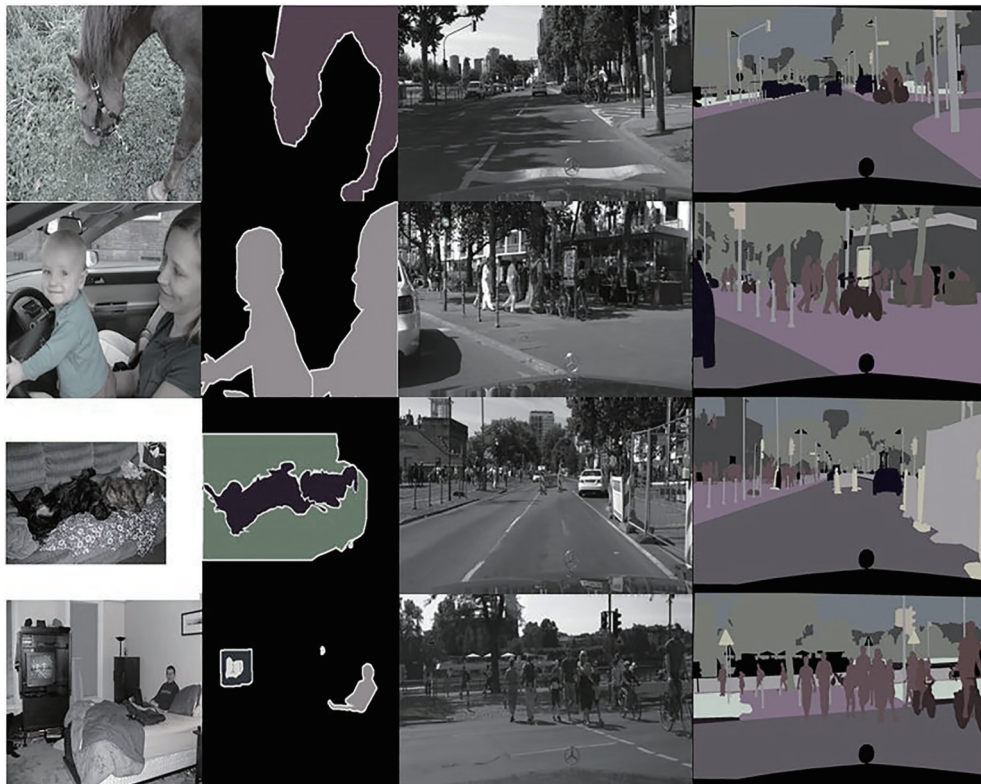


Fig.5 Samples of PASCAL VOC and Cityscapes datasets  
图 5 PASCAL VOC 数据集与 Cityscapes 数据集样例

2.2 评价指标

语义分割中常用交并比(IoU)作为评价指标,其含义为预测的像素点语义标签与标注的像素点语义标签的重合度,以及所有类别的平均交并比(mean IoU, mIoU),定义为:

$$IoU(i) = \frac{p_{ii}}{\sum_j (p_{ij} + p_{ji}) - p_{ii}}; mIoU = \frac{1}{C} \sum_c IoU(c) \tag{8}$$

式中:  $p_{ii}$  表示像素点预测为  $i$  类且标签也为  $i$  类;  $p_{ij}$  表示像素点预测为  $i$  类而标签为  $j$  类;  $p_{ji}$  表示像素预测为  $j$  类而标签为  $i$  类;  $C$  表示数据集的类别数。

2.3 训练过程

实验中使用 Tensorflow 深度学习框架,并用 1 块 NVIDIA TITAN GPU 进行训练。输入图片分辨率为  $500 \times 500$ ,数据批量设置为 2,使用线性下降法训练 10 万个  $step$ 。首先单独对生成模型进行训练至收敛,此时设置初始学习率为  $10^{-4}$ ,并逐步衰减至  $10^{-5}$ 。然后加入对抗模型训练。由于对抗模型此时处于初始化状态,而生成模型已接近收敛,需要使用相对较大的学习率帮助训练对抗模型,重置生成模型的学习率为  $3 \times 10^{-5}$ ,对抗模型的学习率设为  $3 \times 10^{-4}$ 。参数  $\lambda$  设置为  $10^{-2}$ 。在优化算法方面,生成模型使用 Adam 优化器,对抗模型则使用 Momentum 优化器,其中 Momentum 设为 0.9。在实验中,ResNet 与 VGG 模型均预先加载了用 ImageNet 预训练的权重,再使用语义分割数据集进行分割任务的训练。

2.4 实验结果及分析

共设置 4 组实验,分别为: VGG16, ResNet50, VGG16-G 和 ResNet50-G。通过对比分割模型中是否引入生成对抗网络后对图片分割的结果,确定生成对抗网络的效果。在 PASCAL VOC 数据集和 Cityscapes 数据集上的结果如表 3 所示。

表 3 四种分割模型测试结果对比  
Table3 IoU comparison of four segmentation models

datasets networks	IoU/%			
	VGG16	ResNet50	VGG16-G	ResNet50-G
PASCAL VOC	71.62	74.29	73.18	75.46
Cityscapes	68.34	75.85	70.27	77.40

从表 3 中可看出,在 PASCAL VOC 和 Cityscapes 数据集上,引入生成对抗网络后分割精确度均有一定的提升。PASCAL VOCs 数据集中, VGG16-G 的 IoU 提升了 1.56%, ResNet50-G 的 IoU 提升了 1.17; Cityscapes 数据集中, VGG16-G 的 IoU 提升了 1.93%, ResNet50-G 的 IoU 提升了 1.55%。2 个数据集上, ResNet50 的表现均优于 VGG16,这是由于 ResNet50 中引入的空洞卷积扩大了感受野,使深层特征图能够囊括范围更广的信息,且 ResNet50 的残差结构能够很好地补充浅层特征至深层特征中,这对于需要语义特征以及空间特征的语义分割任务是有益的。VGG16 在 Cityscapes 数据集上的表现不如 PASCAL VOC 数据集,是因为 PASCAL VOC 数据集的图片较小,原图尺寸仅为  $500 \times 375$ ,分割对象数量少,尺寸较小,对分割错误的容错性更高一些。而 Cityscapes 数据集中图片尺寸为  $2048 \times 1024$ ,且是对交通场景中所有目标进行分割,分割对象数量多,尺寸差距大,较容易出现小物体缺失、大面积颗粒化现象。因此在 Cityscapes 数据集上 VGG16 的表现明显不如 ResNet50。而通过引入生成对抗网络结构,能够改善上述缺陷,使分割更加精确,并减轻颗粒化现象,如图 6 所示。

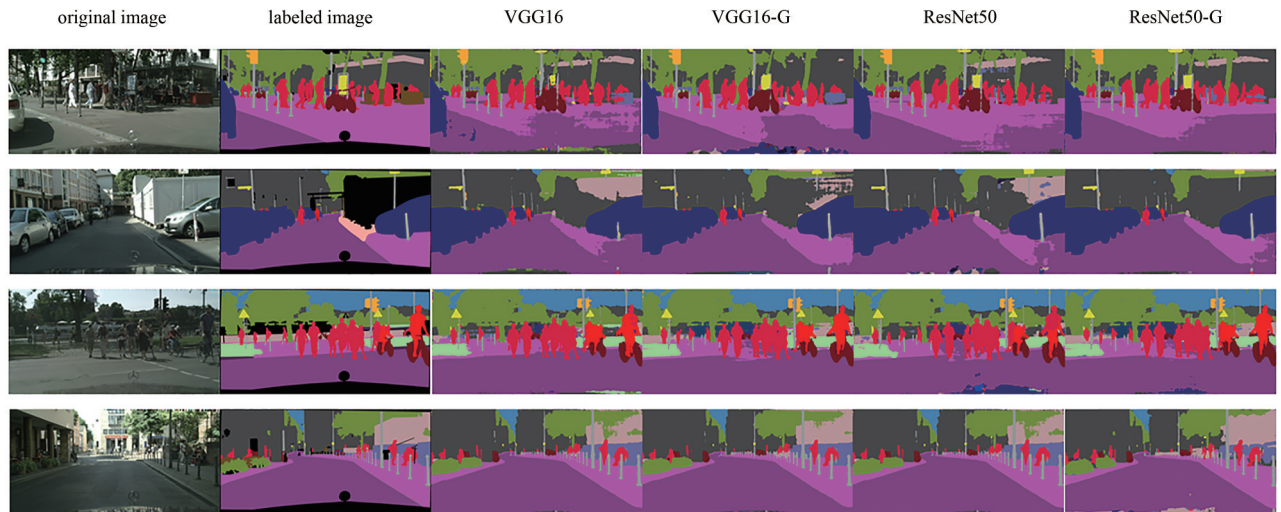


Fig.6 Samples of tests on the Cityscapes dataset  
图 6 Cityscapes 数据集上各模型测试样例

图 6 中, VGG16 在大物体, 如马路上, 较容易出现大面积丢失和颗粒化, 而加入生成对抗网络后, 明显得到改善。ResNet50 出现相同问题, 但程度较轻, 引入生成对抗网络后效果变好。为进一步对比 4 种模型的区别, 在实验中选择 Cityscapes 数据集中的 6 类进行 IoU 对比, 如表 4 所示。

表 4 Cityscapes 数据集上 4 种分割模型测试结果对比  
Table4 IoU comparison of four models on the Cityscapes dataset

network	IoU/%					
	road	traffic lights	plants	sky	pedestrian	vehicle
VGG16	94.73	64.12	86.11	90.46	76.88	90.61
ResNet50	98.39	71.37	92.70	94.98	82.49	95.48
VGG16-G	96.64	64.15	87.35	90.87	77.52	92.08
Res50-G	98.48	73.41	93.02	95.35	83.71	95.73

表 4 中, 在引入生成对抗模型结构后, 各类分割精确度均有一定提升。对马路、天空、车辆等占据图片面积大, 目标数量多的类别, 提升效果相对较好; 对交通灯、行人等面积小, 数量少的类别, 提升效果较弱。这是由于小目标难以被分割到, 在前期过滤中信息容易丢失, 需要通过补充浅层信息等方式来捕获小目标的位置。

### 3 结论

将生成对抗网络这一结构引入语义分割的任务中, 引导现有的语义分割模型捕捉像素间的二元势函数关系, 在分割模型结构、参数量、计算量不变的基础上提升分割效果。相较于传统的使用 CRF 等后处理技术提升模型分割精确度的方式, 具有速度的优势。通过在 VGG16 和 ResNet50 这两个常用的语义分割基础模型上的实验结果, 证明了使用生成对抗网络能有效提升网络的分割精确度。在训练中生成对抗网络训练难度较一般的神经网络更高, 特别是在对对抗模型的训练中, 模型较易崩溃, 需在训练过程中进行调整。后续针对这一问题做进一步的探究, 通过网络结构、损失函数等方面改进, 降低训练难度, 提高利用生成对抗网络提升分割精确度方法的鲁棒性。

#### 参考文献:

- [ 1 ] LONG J,SHELHAMER E,DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014,39(4):640-651. doi:10.1109/TPAMI.2016.2572683.
- [ 2 ] KRIZHEVSKY A,SUTSKEVER I,HINTON G. ImageNet classification with deep Convolutional Neural Networks[J]. Advances in Neural Information Processing Systems, 2012,25(2):1097-1105. doi:10.1145/3065386.
- [ 3 ] SZEGEDY C,LIU W,JIA Y,et al. Going deeper with convolutions[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston,MA:IEEE, 2015:1-9. doi:10.1109/CVPR.2015.7298594.
- [ 4 ] SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J/OL]. arXiv:1409.1556v2, 2014.
- [ 5 ] SILBERMAN N,HOIEM D,KOHLI P,et al. Indoor segmentation and support inference from RGBD images[C]// European Conference on Computer Vision(ECCV). Florence,Italy:[s.n.], 2012:746-760.
- [ 6 ] CHEN L C,PAPANDREOU G,KOKKINOS I,et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. Computer Science, 2014(4):357-361.
- [ 7 ] RONNEBERGER O,FISCHER P,BROX T. U-Net:convolutional networks for biomedical image segmentation[C]// International Conference on Medical Image Computing & Computer-Assisted Intervention. [S.l.]:Springer, 2015:234-241. doi:10.1007/978-3-319-24574-4\_28.
- [ 8 ] GOODFELLOW I J,POUGET-ABADIE J,MIRZA M,et al. Generative adversarial networks[J]. Advances in Neural Information Processing Systems, 2014(3):2672-2680. doi:10.1145/3422622.

#### 作者简介:

刘可心(2000-), 女, 学士, 主要研究方向为图像处理与大数据分析,email:1932129119@qq.com.