
Notes

Notes

Models

Tips

Meeting 20180323

PyTorch 指定使用GPU

ubuntu执行.sh文件的方式和区别

Python pathlib.Path() Examples

unittest

Tiny Face Detector, CVPR 2017

classmethod

Log

2018-3-28

Effective use of dilated convolutions for segmenting small object instance in remote sensing imagery

Abstract

Models

- **ultrasound-nerve-segmentation**是一个经典u-net结构，非常完整。有四个下采样层，不足之处是数据处理部分采用的将数据全部一次读进内存，不适合训练数据很大的task。已经改进。
- **ZF_UNET_224_Pretrained_Model**是一个有预训练模型的u-net，每一个卷积后添加了BatchNormalization层，并提供了完整的demo和batch_generator函数。不足之处是提供的模型输入224*224的，并且只匹配输入图像channel=3的图像。
- **u-net** 其中有关于使用resnet训练的代码，这个值得看一下，在train_res.py中。
- image-segmentation-keras 包括利用VGG和多个模型结合，包括FCN，PSP，Seg，Unet，所以主要学习一下VGG是怎么和这些网络结合的。
- **kaggle_carvana_segmentation** - teraus主要pre-model，loading 已经下载完成，vgg11-bbd30ac9.pth。train_loader has problems。需要先将其他的跑起来，看看其他中的train_loader 有没有可以借鉴的。asanakoy是使用的torch2.0版本，但是机器上安装的是torch0.1
- **Kaggle-Carvana-Image-Masking-Challenge** 已经跑起来，包括128，256，512，1024三种输入大小的u-net，另外还有关于multi_GPU和multithread的程序。
- image-segmentation-keras loading vgg16模型，还没有成功

Tips

- Ubuntu [批处理修改图片大小](#)：imagemagick。

Install

```
sudo apt-get install imagemagick
```

目录里的所有图片都缩小一半

```
$ cd <图片所在路径>
```

```
$ mogrify -resize 50% -format jpg
```

指定大小来缩放图片也是可以的

```
mogrify -resize 800x600 -format jpg *
```

更多应用查看[这里](#)

Meeting 20180323

- DeepLab
- MaskLab
- 空洞卷积
- Aggregation Network for instance segmentation
- DeepLung

PyTorch 指定使用GPU

- 终端设定 `CUDA_VISIBLE_DEVICES=1 python main.py`
- python 代码中设定

```
import os
os.environ["CUDA_VISIBLE_DEVICES"] = "2"
```

ubuntu执行.sh文件的方式和区别

例如以下test.sh脚本：

```
@!/bin/bash(@为#)

read -p "Please input your first name:" firstname
read -p "Please input your last name:" lastname
echo -e "\nYour full name is: $firstname $lastname"
```

- `sh test.sh`
- `bash test.sh`
- 使用点 . 执行 `.test.sh`，但是使用之前必须为文件添加执行权限 `$ chmod +x test.sh`，添加完执行权限之后，便可以使用 `./test.sh` 来执行脚本，该方式与 `bash test.sh` 是一样的，默认使用 `bin/bash` 来执行我们的脚本。
- 使用source执行
使用source则也能够直接执行我们的脚本：`source test.sh`

区别在[这里](#)

Python pathlib.Path() Examples

unittest

Tiny Face Detector, CVPR 2017

classmethod

Log

2018-3-28

- NormalCT已经测完，但是存在一些问题，已经写在.md文件中。明天接着测试其他的。最好算一下精确度之类的。
- 关于keras模型的保存问题，因为保存多了模型，容易混淆，不知道哪个对应哪个，不知道哪个好，哪个差，因此要写log。

Effective use of dilated convolutions for segmenting small object instance in remote sensing imagery

Abstract

由于CNN等网络的发展，高分辨率遥感图像的语义分割取得了很大的进步。但是却忽视了对遥感图像中小并且密集的对象分割。本文提出了一个附加在dilated front-end model 之上的局部特征提取模型LFE。由于核的稀疏性，一味增加膨胀因子会使聚集局部特征失败，并且会损伤小目标。解决方法是降低膨胀因子来聚合局部特征。

1. Intorduction

卫星图像和街道图像的差异，(i).目标大小不同，卫星图像中的目标对象非常小；(ii).目标对象的分布不一样，卫星图像中的目标分布非常密集。解决这种分布密集的小目标，一种方法是使用图像中的上下文信息。【26】是关于这个的讲解。CNN中网络是通过下采样来获得上下文信息的，虽然下采样层能够有效地扩展感受野，但是忽视了另一个重要问题：分辨率。在网络中，下采样层使特征图越来越小，这也就丢失了小目标的细节，即使通过：skip connection 或者 hypercolumns也难以重建。因此我们需要一种在扩大感受野的同时不损伤分辨率的方法。

空洞卷积是一种不错的方法，但是在空洞卷积中核权重的排列是通过膨胀因子来扩展的。增加这个因子，权重就变得稀疏，卷积核的大小也会相应的增加。在每一层单调的增加膨胀因子，感受野就会变大并且不会损伤分辨率。所以空洞卷积在计算机视觉上表现非常不错。

但是单纯的使用空洞卷积并不总是能提高性能。一味地增加膨胀因子，虽然对分辨率和保持上下文信息很重要，但是却会损坏小目标对象。另外现在对于视觉方面研究的趋势就是配备更多的空洞卷积层，然而对于分割小目标物体却要另辟蹊径。

我们提出了一个LFE模型，这个模型由降低膨胀因子的多个卷积层组成。在将这个模型附加在一个增加到空洞卷积之上。这样的组合是非常合适的，因为通过LFE模型，内核的权重变大，这样局部特征就会聚合。

通过多个数据集的评估，本文所提模型性能在小物体 优于现在最好的u-net模型和DeepLap模型。另外还进行了ERF分析(???)。

2. Realted work

语义分割是对每一个像素进行标记。FCN将分类网络扩展成一个可以对像素进行标记的网络，在语义分割领域已经取得了巨大的成就。语义分割的另一个很有挑战性的问题是如何精确定位。简单的扩展分类网络并不能提取清晰地目标边框，这是因为分类网络的特征在空域上非常粗糙(??)。另一种方法就是就是编码-解码结构。e.g. segnet。在编码部分，低分辨率的语义特征首先被提取出来，在解码部分利用编码中最大pooling层所选在的位置为线索，对特征图的空间分辨率进行重建。U-net在解码过程中，没有使用最大pooling层提供的位置信息，而是通过 skipping和组合编码过程中的低层次特征逐步的改善整个网络的特征。第三种方法是利用空洞卷积，空洞卷积是在不损失分辨率的情况下有效(effectively, 最大的(??))的扩大感受野。别人也用到过，但是他们适用于街道图像的语义分割，本文用于遥感图像小目标的实例分割。其他的在小目标分割精度上有提高的，【13】使用了样本平衡损失函数。【16】根据每一个目标与边框之间的距离对输入图像的像素进行分类(???)，这种方法能够得到目标边框的精确位置。

本文的目标是遥感图像语义分割和目标实例的检测。以前通常有两个步骤：提取目标mask和分类。

3. 方法

整体结构：front-end 模块，local Feature extraction 模块，head 模块。front-end模块用于提取包含大的上下文信息的特征，所以膨胀因子是逐渐增加的。后面的lfe模块用于汇聚front-end模块获得分散特征，所以lfe模块有一个膨胀因子逐渐降低的特殊结构。最后的head模块是一个和输入具有相同分辨率的概率图。这个模型就是一个卷积版的全连接分类网络，例如VGG。

front-end 模块

这个模块的主要作用是汇集上下文信息。CNN中一般使用下采样来扩大感受野，但是下采样会降低学习到的特征图的分辨率，必须要使用下采样的原因是为了避免网络参数爆炸。

为了获得大的感受野并且保持高的空间分辨率，我们采用空洞卷积。空洞卷积能在扩大感受野的同时保持分辨率不变。空洞卷积使用稀疏的并且均衡化的权重作为特定的核。核的大小和稀疏权值的区间随着膨胀因子呈指数增长。通过增加膨胀因子，感受野也会随着扩大的核变大。文中将该模块的所需下采样层都换成了空洞卷积。虽然这对于小的建筑物很有效，但是在空洞内核方面有两个问题需要考虑。

LFE模块

空洞卷积会导致两个问题：（1）相邻元素之间的空间一致性变差（2）无法再高层中提取局部结构（??）。

假设进行卷积核为2的一维卷积，膨胀因子是2。左边顶层的蓝色单元只会受低层蓝色单元影响。这些相同颜色的单元组成了信息金字塔，同时也确定了顶层单元能够看到的区域。由于膨胀因子为2，所以信息金字塔两个相邻单元之间没有重叠。这种相邻单元之间的不一致性，导致了最终的输出存在锯齿现象。

局部特征提取问题

信息金字塔确定了低层单元的影响区域。同时，信息金字塔在低层的两个相邻单元之间不会重叠。顶层的所有单元接收到的信息只能来自这两个中的一个，不能同时接受两个。这就意味着顶层的所有单元都无法获悉低层两个单元之间的结构信息。

在LFE模块中，随着膨胀因子 r 的增加。为了从目标的特征中识别局部结构信息，若是目标足够大，那么这不是问题。这种情况下，局部结构信息能够通过低层的密集核提取出来。对于小目标区域，一些局部结构信息需要在高层被提取出来，这是因为需要通过大的上下文信息来识别它们。但是，随着膨胀因子的增加，在无重叠的信息金字塔中无法提取局部结构信息。

局部特征提取模块

为解决这两个问题，本文提出了降低膨胀因子的局部特征提取模块。思想是我们在增加膨胀因子的后面附加一个降低膨胀因子的结构，邻域单元之间的信息金字塔就会存在连接。因此，降低膨胀因子的结构逐渐恢复了邻域单元之间的一致性，并提取了高层的局部结构信息。试验中显示LFE模块对小目标很有用。