

VGGnet论文总结 (VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION)

VGGNet的主要贡献:

- 1、增加了网络结构的深度
- 2、使用了更小的filter (3*3)

1 introduction

这部分主要说明了, 由于在所有的卷积网络上使用了3*3的filter, 所以使整体网络的深度加深。最后在ILSVRC取得的成绩也是十分明显的。

2 Convent Configuration

2.1 ARCHITECTURE

在前边的卷积层上, 使用3*3的filter, 有时也使用1*1的filter (相当于在输入channels上进行线性变换)。卷积操作的步长是1。总共有5个max-pooling层, 大小是2*2, 步长设定为2, 跟在一些卷积层之后。在所有卷积层之后, 跟着3个全连接层, 前两个全连接层有4096个channels, 最后一个全连接层有1000个channels对应1000种分类。最后一层是soft-max层。

所有的隐藏层都有整流函数(ReLU)。

2.2 CONFIGURATIONS

配置了A-E五种结构的网络, 深度分别对应11,11,13,16,19。卷积层的宽度从64开始, 每经过一个max-pooling层, 大小乘2, 直到512。

2.3 DISCUSSION

- 1、VGGnet使用了3个非线性整流层, 相比于只使用单个, 决策函数更具有判别力。
- 2、使用3层叠加的3*3的filter的卷积层, 要比使用一个7*7的filter的卷积层的参数要减少81%。

另外, 使用1*1的filter的卷积层可以增加决策函数的非线性特征。

3 CLASSIFICATION FRAMEWORK

3.1 TRAINING

用momentum方法的mini-batch方法, batch的大小为256, momentum设为0.9。weight decay采用L2正则化方法, 惩罚因子为0.0005。对于全连接层的前两层采用dropout的正则化方法, dropout的大小设为0.5。学习率设为0.01。

对A这样深度不算深的网络来说, 使用随机初始化去训练参数, 然后用A训练得到的参数去初始化其他更深的网络。

用来训练的图片的规格大小: 第一步设定要训练的图片的规格S = 256, 当训练规格S=384的网络时, 先用S=256的结果初始化, 然后使用0.001的训练速率。第二步, 在一定的范围内随机截取要训练的图片, 然后用S=384训练得到的结果去初始化这个网络。

3.2 TESTING

在测试过程中, 使用了两种方法, 第一种是dense evaluation, 使用FCN。第二种是multi-critevaluation, 通过截取获得一个大的数据集, 能是提高结果的精确度。

3.3 IMPLEMENTATION DETAILS

实现基于caffe, 做了一定的修改, 允许在单操作系统的多核GPU上进行训练。用多GPU进行并行计算每个batch的梯度, 当所有的GPU都计算完成之后, 求所有batch得到梯度的平均值。

4 CLASSIFICATION EXPERIMENTS

4.1 SINGLE SCALE EVALUATION

- 1.LRN不能降低错误率, 后续网络中不再使用

2.随着网络层数的加深，分类错误率逐渐降低

3.scale jittering可以使表现结果更好

4.2 MULTI-SCALE EVALUATION

在训练时用scale jittering可以使结果表现的更好，比使用单一规模

4.3 MULTI-CROP EVALUATION

单一使用multi-crop evaluation 要比单一使用dense evaluation效果好，两个方法同时使用时，要比单一使用任意都好。

4.4 CONVNET FUSION

将几个模型的soft-max分类策略的输出求平均后再用于识别，这样可以提高最后的表现。

5 CONCLUSION

通过在大规模的图片分类上评估深度卷积神经网络，表明深度有益于分类的精确度，并且通过使用加深了的传统卷积神经网络，能达到在ImageNet数据集上的最好表现。

附录A LOCALISATION

A.1 LOCALISATION CONVNET

与分类的区别是，在最后一个全连接层，使用bounding box代替 class scores。如果bounding box在所有类交叉共享的，那么最后一层就是4-D，如果是特定类，那最后一层就是4000-D。

Training. 和分类主要的不同是，用Euclidean loss替换logistic regression objective，在惩罚bounding box预测偏差的时候。训练模型时，不使用scale jittering

Testing. 测试使用两种不同的测试方法，

- 1、bounding box只在图像裁剪中心得到，用于比较在验证集下不同的网络限制，bounding box的预测效果。
- 2、在整个图像上，密集的应用定位网络。与分类任务不同的是，用一组bounding box的预测代替class score map。为了提出最后的结果，使用贪婪融合过程，先融合相近的预测，然后用class scores进行估计。

A.2 LOCALISATION EXPERIMENTS

Settings comparison. 1、使用PCR的效果要比使用SCR的效果好。2、微调所有层比仅微调全连接层的效果好

Fully-fledged evaluation 使用最优的设置（PCR, fine-tuning of all layers），图像的scale对结果有影响。

Comparison with the state of the art.

附录B GENERALISATION OF VERY DEEP FEATURES

标签: 论文



0

0

[+加关注](#)

« 上一篇: [内核笔记](#)

» 下一篇: [配置caffe matlab 中遇到的坑](#)

posted @ 2017-09-12 17:05 Unicoe 阅读(279) 评论(0) 编辑 收藏