关于upsampling的相关知识在cs231n的lecture11

Fully Convolutional Networks for Semantic Segmentation

Abstract

Introduction

Related work

Fully Convolutional Networks

Segmentation Architecture

Results

Conclusion

Abstract

卷积网络是强大的视觉模型，可以产生特征层次结构。本文展示的网络是端到端。像素到像素的，表现优异。本文构建了全卷积网络，该网络采用任意大小的输入，并通过有效的推理和学习产生相应大小的输出。本文定义了一种跳跃结构，该架构将来自深层粗糙层的语义信息与来自浅层细节层的外观信息相结合，以产生准确和详细的分割。网络处理时间也很快

卷积网络推动了识别的进步。Convnet不仅改善了整体图像分类，而且还在结构化输出的局部任务上取得进展。这些包括边界框对象检测，部分和关键点预测。从粗略推理到细微推理的下一步是对每个像素进行预测。先前的方法使用了用于语义分割的卷积，但是具有缺点。

全卷积网络在语义分割上训练端到端的像素到像素，超过了现有技术。据我们所知，这是首次训练FCN端到端（1）进行像素预测，（2）进行有监督的预训练。

语义分割面临语义和位置之间的内在问题：全局信息解决了本地信息在何处解决的问题。

深度特征层次结构将位置和语义信息编码进一个局部到全局的金字塔内。

我们定义了一个跳跃架构（skip architecture）来利用这个特征谱，它结合了深层的，粗糙的语义信息和浅层的，细微的外观信息

为什么CNN需要对输入图像尺寸做要求：

因为CNN最后有全连接层，而全连接层的输入是固定大小的。