第二篇论文问题解决

由于前期理解不到位，误以为一些没给出明确值的参数是需要计算的，其实这些都是可以进行调节的参数，解决了这一问题后，对于第二篇论文的问题进行了精简，并搜索相关资料，查阅了相关论文，总结回答如下：

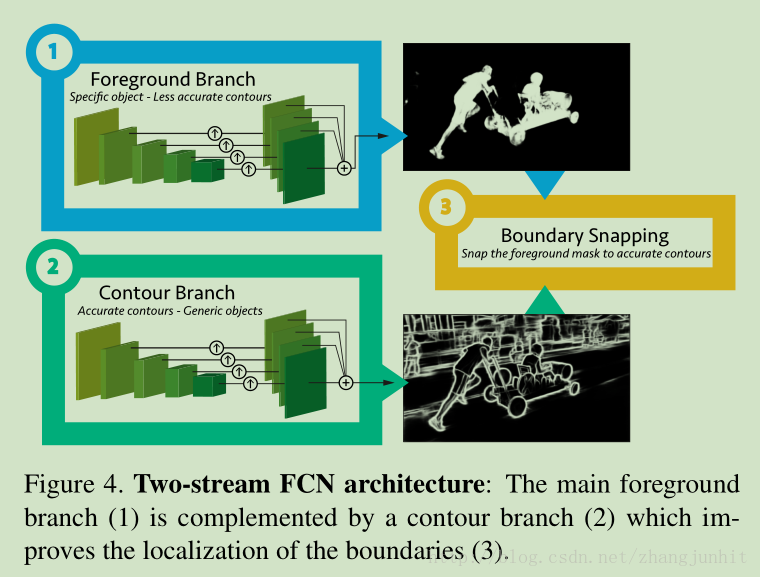
1. Part3.2:
2. CNN将120key-frames 分为前景背景的算法？

**One-Shot Video Object Segmentation**（CVPR2017），基于单帧标记的视频物体分割，对于一个视频中的某一个物体，一个基于全卷积神经网络的结构。这个结构在ImageNet进行预学习，对前景进行语义分割，最终利用一张标注的图片来实现整个视频序列中该物体的分割。所有帧都是独立处理的，但是结果显示连贯稳定。

基于CNN的One-Shot Video Object Segmentation (OSVOS) 算法总体的思路如下：

先在 ImageNet 训练一个图像分类模型，得到的网络称之为 Base Network，接着我们 在 DAVIS 训练一个分割网络，得到 Parent Network，最后我们在目标视频上微调，得到 Test Network。前两个网络的训练都是 offline， 最后一个网络的训练是 online。这个online 时间越长效果越好。

**对于单张的训练：an object->this particular object**



**End-to-end trainable foreground FCN**

这里我们采用了一个 **语义分割网络**，参考文献【30】，基于 VGG模型，在准确定位分割预测上做了修改。就是上图的 **foreground branch**，对于这个分割网络的损失函数，我们参考文献【51】的 **pixel-wise cross-entropy loss**，同时了采取了文献【51】解决 **imbalance between the two binary classes** 的方法。在微调部分将用于分类的全连接层移除，使用了image-to-image inference。VGG的网络结构中5个stages 包含了一组ReLU层。在stage之间，池操作缩小了feature maps的规模。

这个前景分割主要是判断物体的有无和大致位置，不能给予精确的边界位置信息，这里我们又加入了 contour 信息， 边缘检测CNN网络的训练完全是离线的， train the contour branch only offline

有了边缘信息，我们使用这些边缘信息来得到精确的物体分割， the use of the **Fast Bilateral Solver (FBS)** [2] to snap the background prediction to the image edges

1. Part 3.3

1. Hc？shading of a vertex with spherical harmonic components c.：

具有球面调和分量c的**高通滤波函数**

1. Part3.4
2. 关于关键帧的选取以及十个语义分割：

**Deep human parsing with active template regression**

论文中详细解释了人体语义分割问题。论文中的label数量选取并没特殊提到10的问题，而是用了更多的17或18（包括眼镜、围巾等配饰）。猜测应该是考虑到原论文人物的衣着特点进行的选择，认定10个label在运用到论文2时最佳。