Unified framework for weak body parsing

1. 输入输出

输入为头部及躯干的candidates，输出为最终configuration

1. Unified体现在哪里

求解最终configuration时，联合了part信息以及衣服信息

1. 为什么需要unified
   1. 提高识别准确率。part与衣服能互相增强，互相补充，如果分为两步，先做part，在分割衣服，那仅仅有第二步的结果会依赖第一步，第一步挂了第二步也跟着挂，所以需要在inference的时候把两类信息同时加进来
   2. 降低计算代价。一方面，从search的角度，有实时性需求，framework的输出需要快速完成；另一方面，从衣服属性抽取的角度，属性主要集中于躯干，袖子提供的衣服信息很少。所以weak body parsing完全能handle我们的task。

问题1：怎么把衣服信息加到feature中

Unified framework for clothing attributes?

输入不变，直接输出想要的每个attribute的feature

怎么定义输出结构？

目前attribute包含：领口、logo、纹理

纹理跟躯干的parse等价，

领口可以用bounding box表示，logo用0、1表示

加上头部、躯干，总共3个bounding box，可以用一个13维向量表示

问题1：领口的bounding box需要自适应的大小，意味着搜索空间很大

问题2：怎么把结构信息encode到feature中去。主要是领口的信息怎么跟头、躯干发生关联，领口的bounding box的大小，怎么在feature层面上说明某个大小是最优的

Generalization:

Unified的套路能解怎样一类问题？

e.g. 直接在图片中定位椅子很困难，但椅子往往跟桌子是摆在一起的，两者的信息能相互增强，其结构信息可以用两者的空间距离表征

1.耦合问题

2.要解的是其中一个问题到部分解

3.搜索场景

4.