

结果

算了 100 张图，数据集还是 COCO MVal。

ps. 把所有 Attention 全部存下来不太现实，存 4 张图的所有 global attention 就要吃掉 48GB 的空间了。所以最后还是在线计算的，也不会很慢。

就结果来看，边缘 patch 受到的关注度可能略低于一半，而且层越靠后越低。

边缘 patch 的判断

边缘提取目前用的是 Canny 算子，只要一个 patch 里面有包含边缘，就认为是边缘 patch。所以一张图有可能大部分 patch 都是边缘 patch。手调了一下 Canny 算子的参数，并且喂给 Canny 之前先简单高斯模糊了一下，对大部分图来说边缘提取结果还行。

Overlap Size

attention 中的每行的 top-100 与边缘 patch 做交集，取交集的大小，除以 100。

对于同一层的同一个 head，在每张图片、attention 每行取平均

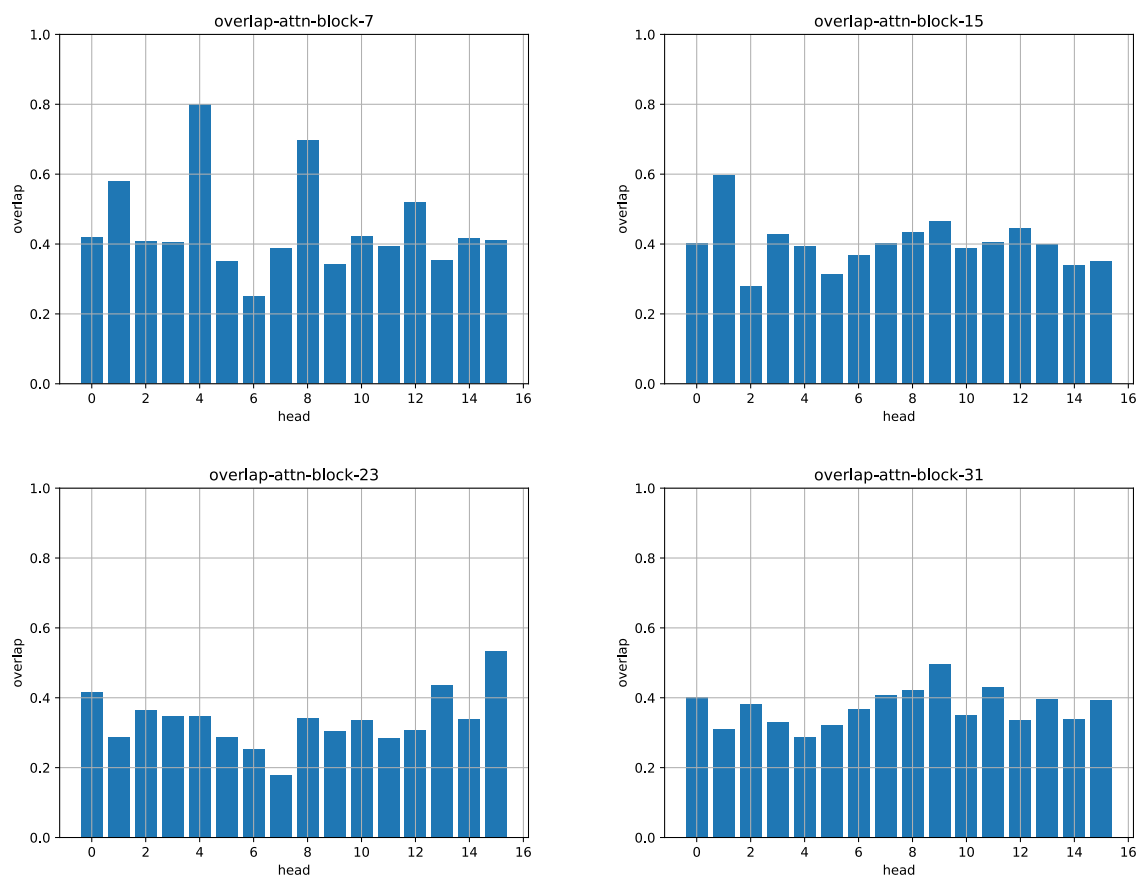


Figure 1: overlap size

Overlap Attention Sum

attention 中的每行的 top-100 与边缘 patch 做交集，算交集部分的 attention 和。

对于同一层的同一个 head，在每张图片、attention 每行取平均

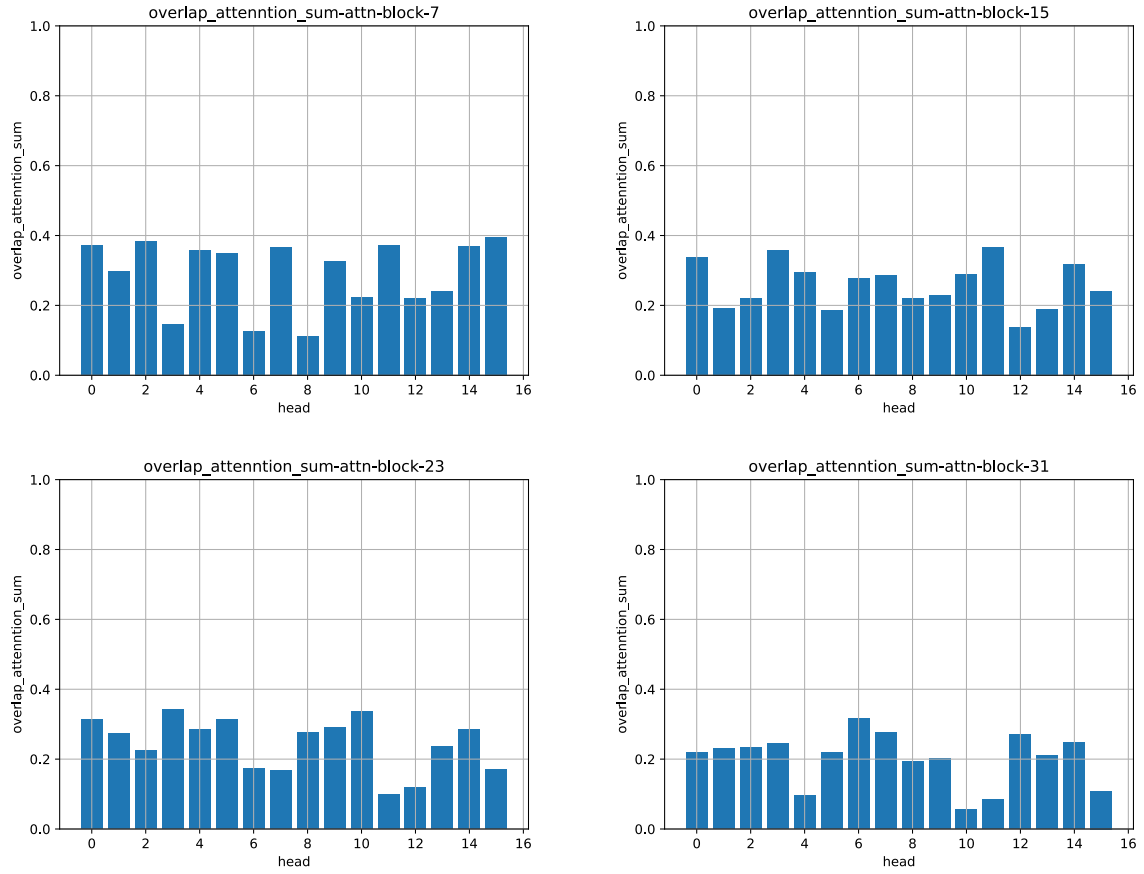


Figure 2: overlap attention sum

Border Attention Sum

attention 中的每行，算边缘 patch 的 attention 和

对于同一层的同一个 head，在每张图片、attention 每行取平均

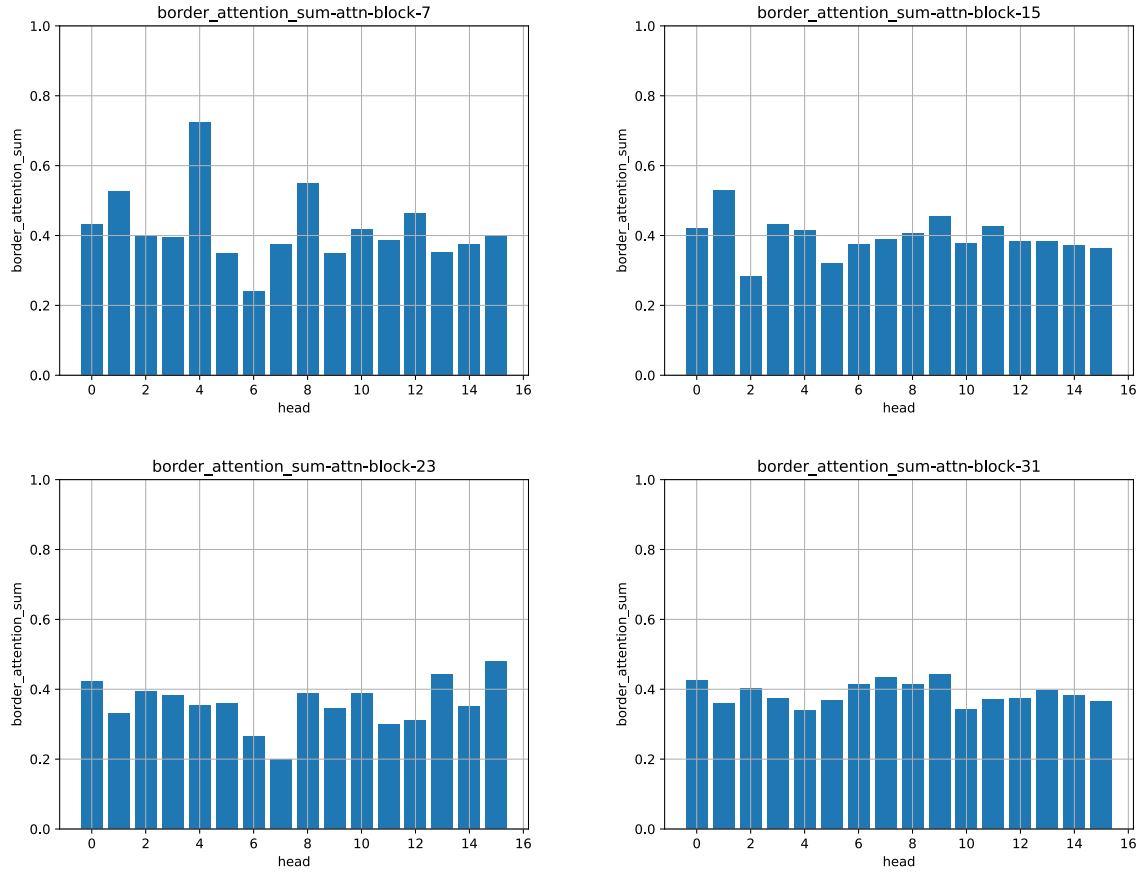


Figure 3: overlap attention sum