**学习周报**

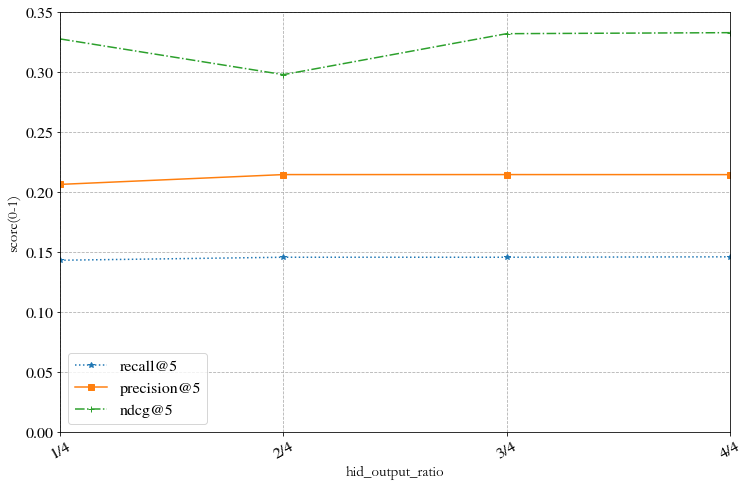
1. **学习内容**
2. 上次论文是精读的，但是没有细讲，这篇论文过程比较复杂

|  |  |  |
| --- | --- | --- |
| 论文 | Spatio-temporal Graph Neural Network based Mask Reconstruction for Video Object Segmentation  利用时空图神经网络来重构掩模，以解决半监督环境下类不可知的视频对象分割 VOS问题（Video Object Segmentation） | |
| 涉及领域 | 目标检测；视频对象分割；空间和时域图 | |
| 主要贡献 | 本文提出了在VOS任务中首次结合时间和空间相关性的时空图神经网络（STG-Net）， 以此重建更准确的视频对象分割掩码。  （1）传统基于检测的方法利用贪婪策略从候选bbox中选择最优，而STG-Net中的空间图网络利用权值边策略考虑所有的候选bbox, 不会丢失局部块的细节  （2）将空间图中选择的掩码按时间排序得到时间图，每个时间节点t的视频帧掩模都会结合前面的历史帧进行改进 | |
| 论文提出的  模型  STG-Net | Proposal  Generation | 本模块目的是离线产生最可能的候选bbox以及对应的mask，主要过程为：  （1）利用**离线且改进的Mask R-CNN算法**针对类不可知的对象生成候选的bbox,；  （2）然后利用检测**置信度detection probability和非极大值抑制NMS**来进一步筛选bbox；  （3）最后利用**离线的deeplab v3+算法**（以ResNet101作为backbone)进行分割并生成mask |
| Spatial Graph Construction | 本模块的目的是利用上一步得到的**候选bbox及mask来构造空间图，并在图上进行聚合更新**。主要过程为：   1. **构建空间图的准备工作**   这里所指的空间图构造是为每一类对象构建一个空间图，而当前帧中候选bbox框选出来的对象是类不可知的，所以我们需要对其进行分类。  这里我们借用**基于光学流的先验知识**来预测当前帧中每个bbox框选对象的类别，具体过程为：  【**获取所有类对象在本帧中的预测位置p】**：假定视频第一帧图像中的bbox框选的类别是已知的，由于视频帧的时间间隔很短所以前后帧间的物体运动可以看做是非常平稳的，于是我们可以基于历史帧某类对象的中心点和bbox位置来预测出此类对象在本帧中的概率学位置p      【**针对选定的bbox预测其类别】**：  对选定的bbox proposal和所有对象在本帧中的预测框计算Iou, 其中得分最高的预测框类别作为选定bbox的类别  【**获取对象在整张图上的完整掩模】**：  得到所有候选bbox的对象类别后，将待研究对象的所有候选bbox对应的掩模放置到完整的图像中得到对应的完整掩模   1. **构造空间图**   使用待研究类别的所有完整掩模构造一个全连接的  空间图，图上边的权重由掩模之间的IoU以及掩码经一层CNN提取的特征间的cosine交互得分两部分控制；     1. **空间图的更新**   空间图上每个节点的隐层表示由一维展平的节点掩模矩阵来初始化，然后经过加权的邻居消息和自身进行聚合，反复迭代多层得到最终更新后的节点表示       1. **重构掩模**   按照给定阈值将节点的最终隐层表示二值化，最后进行二维展开得到重构的掩模 |
| Temporal Graph Construction | 本模块主要是为了利用视频帧的掩模序列构建时间图，然后**利用历史掩模改善当前帧的掩模**。主要过程包括：   1. **掩模选择**   从空间图中所有节点重构后的掩模中利用得分规则选择出最优的掩模。得分由掩模对应的bbox和预测bbox间的Iou和 掩模本身和预测bbox对应掩模的Iou这两个部分组成。     1. **构建时间图**   将每个视频帧图片依据掩模对应的bbox进行裁剪，得到裁剪后的图片和掩模，其中裁剪的图片还需要经过一层CNN进行特征处理；每帧对应的裁剪后图片的特征、裁剪的掩模按照时间顺序构建成一个时间图。  利用时间图来改善当前帧的掩模输出，当前帧掩模的每个像素与前面所有帧掩模的所有像素点都有关系。 |

（2）完成毕设超参数实验和可视化、写毕设论文

按照自己的思路重新设计了超参实验的过程，思路和过程详见论文；

论文初稿大体完成，摘要和部分实验结果还没写上去



1. 展示系统

在线预测模块还没有写完

1. **学习问题**
2. 关于模型运作过程，论文写作比较麻烦
3. 论文篇幅很大
4. **学习计划**
5. 完成毕业论文
6. 完成系统在线预测模块
7. 完成系统需求等其他文档