哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY

2022 年秋季学期

《视听觉信息理解》课程

视觉部分实验

姓名	姚舜宇
学号	1190202107
班号	1903602

1. 算法描述

1. Cluster-NMS：红框部分同 Fast-NMS，黄框部分是 cluster-NMS 特有的部分。该部分是为了避免 Fast-NMS 中存在的删除框时多删的问题。通过该循环，改变了判断一个框是否应该要删除的阈值。

```
def cc_cluster_nms(self, boxes, masks, scores, iou_threshold: float = 0.5, top_k: int = 200):
    # Collapse all the classes into 1
    scores, classes = scores.max(dim=0)
    _, idx = scores.sort(0, descending=True)
    idx = idx[:top_k]
    boxes_idx = boxes[idx]
    iou = jaccard(boxes_idx, boxes_idx).triu_(diagonal=1)
    B = iou
    # 该循环改变了fast-nms中简单的取maxA的方式，可以防止fast-nms中多删除box
    for i in range(200):
        A = B
        maxA, _ = torch.max(A, dim=0)
        E = (maxA <= iou_threshold).float().unsqueeze(1).expand_as(A)
        B = iou.mul(E)
        if A.equal(B) == True:
            break
    idx_out = idx[maxA <= iou_threshold]
    return boxes[idx_out], masks[idx_out], classes[idx_out], scores[idx_out]
```

2. Cluster-NMS+DIoU-NMS：该方法在 cluster-NMS 基础上只做了一个改变：在计算两个 box 的重叠时，不简单考虑两个 box 的重叠，而是将中心位置考虑进来，添加了 DIoU loss 的惩罚项。

```
def cc_cluster_diounms(self, boxes, masks, scores, iou_threshold: float = 0.5, top_k: int = 200):
    # Collapse all the classes into 1
    scores, classes = scores.max(dim=0)
    _, idx = scores.sort(0, descending=True)
    idx = idx[:top_k]
    boxes_idx = boxes[idx]
    iou = diou(boxes_idx, boxes_idx).triu_(diagonal=1)  # 计算iou时考虑中心之间的距离
    B = iou
    for i in range(200):
        A = B
        maxA, _ = torch.max(A, dim=0)
        E = (maxA <= iou_threshold).float().unsqueeze(1).expand_as(A)
        B = iou.mul(E)
        if A.equal(B) == True:
            break
    idx_out = idx[maxA <= iou_threshold]
    return boxes[idx_out], masks[idx_out], classes[idx_out], scores[idx_out]
```

3. Cluster-NMS-Weighted-NMS：该方法首先利用 score 更新 IoU 矩阵，再根据更新后的 IoU 矩阵重新计算框的坐标。

```
D = distance(boxes, boxes)
X = (B >= 0).float()
scores = torch.prod(torch.min(torch.exp(-B ** 2 / 0.2) + D * ((B > 0).float()), X), 0) * scores
idx_out = scores > 0.01
weights = (B * (B > 0.8).float() + torch.eye(n).cuda()) * (scores.reshape((1, n)))
xx1 = boxes[:, 0].expand(n, n)
yy1 = boxes[:, 1].expand(n, n)
xx2 = boxes[:, 2].expand(n, n)
yy2 = boxes[:, 3].expand(n, n)

weightsum = weights.sum(dim=1)
xx1 = (xx1 * weights).sum(dim=1) / (weightsum)
yy1 = (yy1 * weights).sum(dim=1) / (weightsum)
xx2 = (xx2 * weights).sum(dim=1) / (weightsum)
yy2 = (yy2 * weights).sum(dim=1) / (weightsum)
boxes = torch.stack([xx1, yy1, xx2, yy2], 1)
return boxes[idx_out], masks[idx_out], classes[idx_out], scores[idx_out]
```

4. Cluster-NMS+Score Penalty：该方法将 IoU 矩阵每列元素连乘作为惩罚项进行加权。

```
for i in range(200):
    A = B
    maxA, _ = torch.max(A, dim=0)
    E = (maxA <= iou_threshold).float().unsqueeze(1).expand_as(A)
    B = iou.mul(E)
    if A.equal(B) == True:
        break
scores = torch.prod(torch.exp(-B ** 2 / 0.2), 0) * scores # 将得分进行加权
idx_out = scores > 0.01
return boxes[idx_out], masks[idx_out], classes[idx_out], scores[idx_out]
```

2. 实验操作

下载相应代码
配置环境：conda env create -f environment.yml
准备 coco 数据集，按照 README.md 中的指导运行 sh data/scripts/COCO.sh
准备模型
执行 eval.py 文件，选择相应 NMS 方法

3. 实验结果

```
fast_nms with cross class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.41s)
creating index...
index created!
Loading model... Done.

Processing Images [########################################] 4952 /  4952 (100.00%)     7.42 fps
Saving data...
Calculating mAP...

      | all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
------+------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
  box | 30.51 | 49.01 | 47.07 | 44.51 | 41.40 | 37.77 | 32.81 | 26.37 | 17.61 |  7.72 |  0.84 |
 mask | 28.51 | 45.63 | 43.49 | 41.05 | 38.00 | 34.33 | 30.11 | 24.53 | 17.46 |  8.89 |  1.58 |
------+------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
fast_nms without corss class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.37s)
creating index...
index created!
Loading model... Done.

Processing Images                                4952 /   4952 (100.00%)    11.75 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 32.03 | 52.05 | 49.84 | 46.97 | 43.54 | 39.61 | 34.21 | 27.28 | 18.11 |  7.86 |  0.85 |
  mask | 29.69 | 47.94 | 45.62 | 42.93 | 39.66 | 35.69 | 31.21 | 25.29 | 17.92 |  9.09 |  1.60 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
cluster_nms with cross class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.38s)
creating index...
index created!
Loading model... Done.

Processing Images                                4952 /   4952 (100.00%)     7.27 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 31.13 | 50.34 | 48.41 | 45.80 | 42.44 | 38.44 | 33.18 | 26.50 | 17.64 |  7.73 |  0.84 |
  mask | 28.85 | 46.30 | 44.14 | 41.65 | 38.55 | 34.74 | 30.40 | 24.69 | 17.52 |  8.91 |  1.58 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
cluster_nms without corss class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.37s)
creating index...
index created!
Loading model... Done.

Processing Images                                4952 /   4952 (100.00%)    11.31 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 32.37 | 52.63 | 50.59 | 47.85 | 44.24 | 39.99 | 34.34 | 27.29 | 18.08 |  7.85 |  0.85 |
  mask | 29.78 | 48.04 | 45.77 | 43.10 | 39.88 | 35.84 | 31.30 | 25.31 | 17.91 |  9.07 |  1.60 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
spm with cross class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.44s)
creating index...
index created!
Loading model... Done.

Processing Images                                4952 /   4952 (100.00%)     6.90 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 31.92 | 50.78 | 48.90 | 46.44 | 43.39 | 39.68 | 34.49 | 27.74 | 18.59 |  8.28 |  0.94 |
  mask | 29.40 | 47.01 | 44.83 | 42.37 | 39.26 | 35.47 | 31.08 | 25.30 | 17.98 |  9.10 |  1.64 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
spm without corss class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.36s)
creating index...
index created!
Loading model... Done.

Processing Images  [                              ]  4952 /  4952 (100.00%)    10.28 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 33.00 | 53.13 | 51.02 | 48.25 | 44.94 | 40.95 | 35.42 | 28.29 | 18.84 |  8.24 |  0.92 |
  mask | 30.40 | 49.00 | 46.63 | 43.93 | 40.62 | 36.64 | 32.02 | 25.91 | 18.33 |  9.26 |  1.64 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
spm_dist with cross class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.37s)
creating index...
index created!
Loading model... Done.

Processing Images  [                              ]  4952 /  4952 (100.00%)     7.02 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 32.09 | 50.90 | 49.05 | 46.59 | 43.59 | 39.91 | 34.76 | 28.00 | 18.79 |  8.38 |  0.95 |
  mask | 29.50 | 47.06 | 44.93 | 42.45 | 39.36 | 35.61 | 31.26 | 25.47 | 18.07 |  9.19 |  1.63 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
spm_dist without corss class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.39s)
creating index...
index created!
Loading model... Done.

Processing Images  [                              ]  4952 /  4952 (100.00%)     9.49 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 32.98 | 52.96 | 50.89 | 48.18 | 44.90 | 40.96 | 35.46 | 28.34 | 18.88 |  8.26 |  0.92 |
  mask | 30.29 | 48.73 | 46.41 | 43.74 | 40.45 | 36.52 | 31.94 | 25.88 | 18.29 |  9.27 |  1.63 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
spm_dist_weighted with cross class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.39s)
creating index...
index created!
Loading model... Done.

Processing Images  [                              ]  4952 /  4952 (100.00%)     6.94 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 32.29 | 50.91 | 49.06 | 46.67 | 43.68 | 39.97 | 34.92 | 28.33 | 19.29 |  8.96 |  1.08 |
  mask | 29.52 | 47.05 | 44.99 | 42.45 | 39.38 | 35.63 | 31.24 | 25.44 | 18.11 |  9.29 |  1.64 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```

```
spm_dist_weighted without corss class
Multiple GPUs detected! Turning off JIT.
Config not specified. Parsed yolact_base_config from the file name.

loading annotations into memory...
Done (t=0.38s)
creating index...
index created!
Loading model... Done.

Processing Images [                              ]   4952 /   4952 (100.00%)     9.65 fps
Saving data...
Calculating mAP...

       |  all  |  .50  |  .55  |  .60  |  .65  |  .70  |  .75  |  .80  |  .85  |  .90  |  .95  |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
   box | 33.45 | 53.01 | 50.84 | 48.25 | 45.07 | 41.04 | 36.10 | 29.32 | 20.30 |  9.38 |  1.21 |
  mask | 30.35 | 48.76 | 46.51 | 43.74 | 40.43 | 36.69 | 31.89 | 25.94 | 18.39 |  9.48 |  1.69 |
-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+
```