

# 中国农业人工智能创新创业大赛 方案分享

2021 0722

**陈昱翔**

- MgeEditing代码库介绍
- 比赛方案介绍
- 一些代码细节
- 未来展望

# 1.MgeEditing代码库介绍

# 为啥有这样一个库？



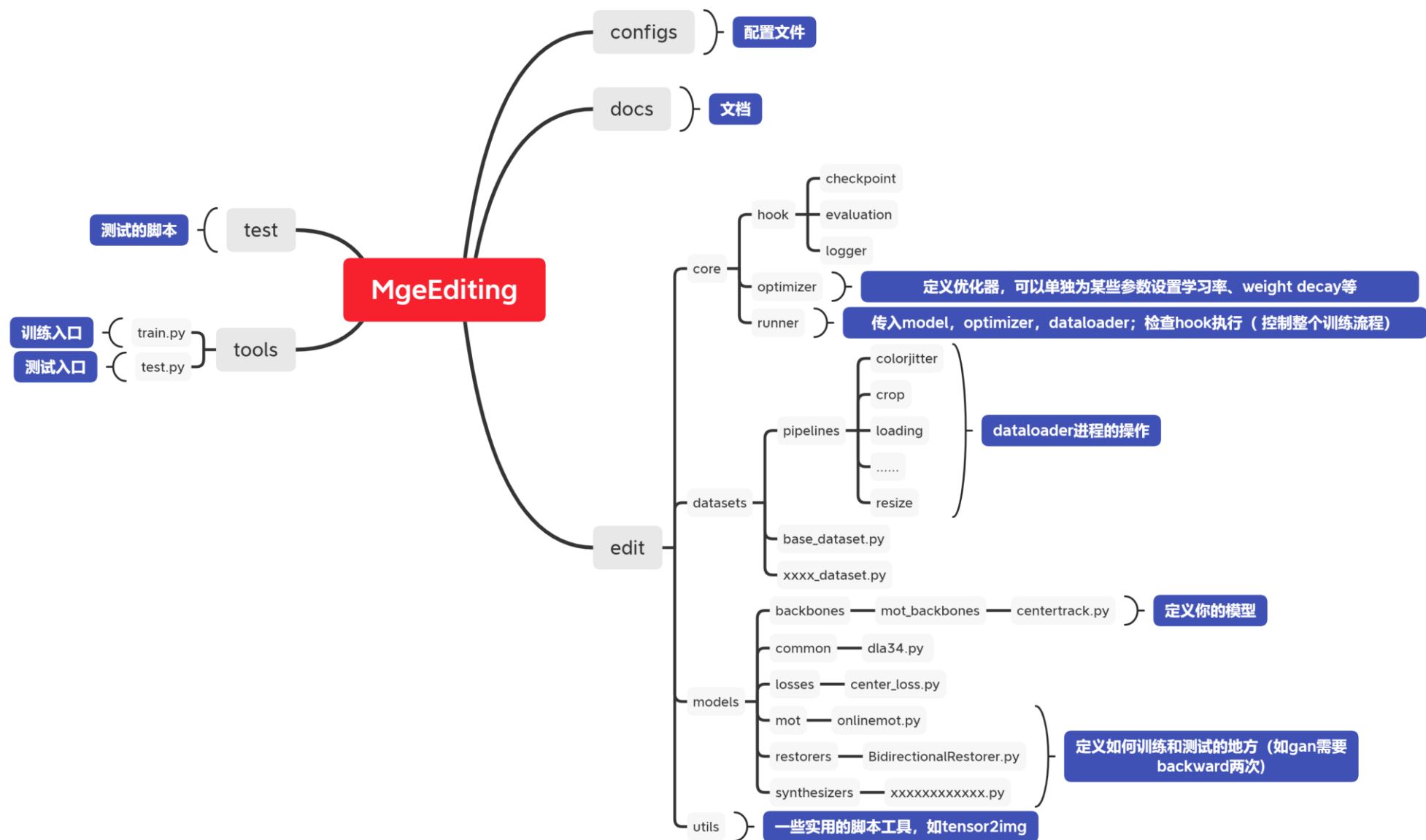
奖金池: ¥ 87000

旷视人工智能开源大赛

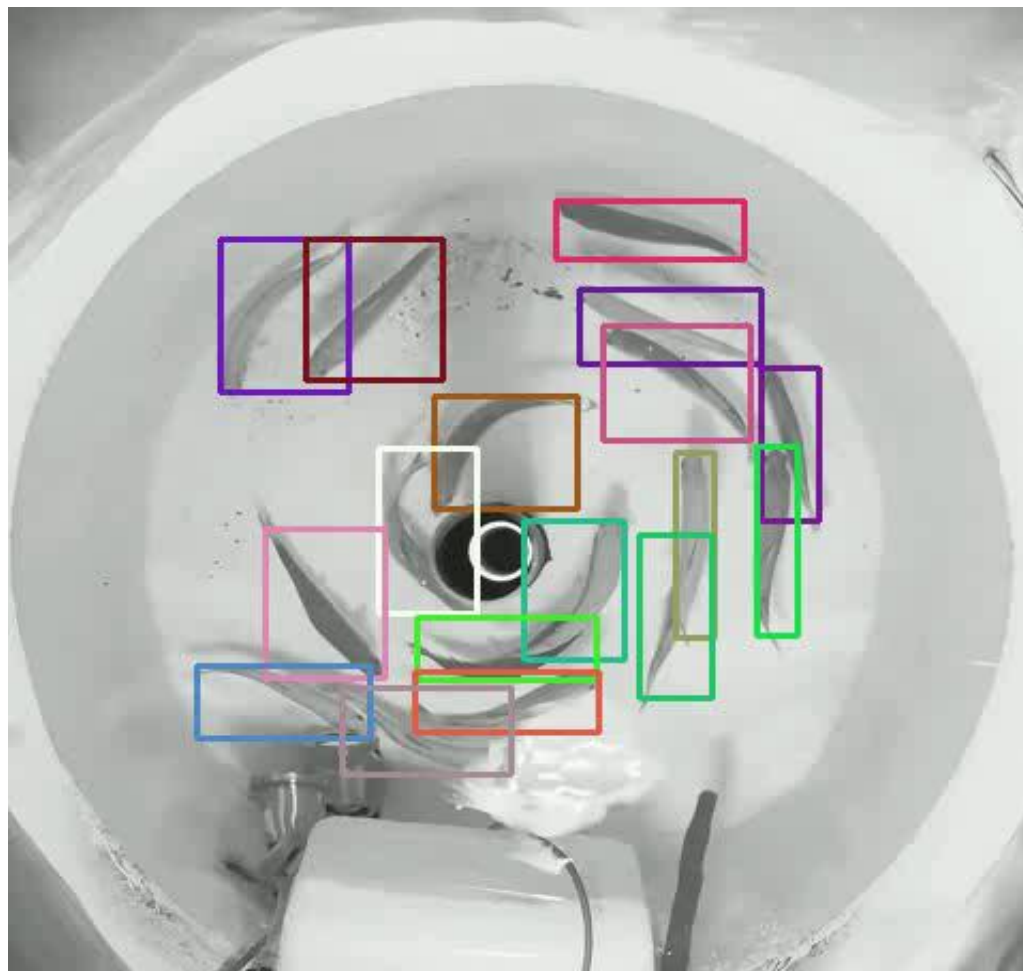
已结束 | 视频超分辨率 | 190 个人已参赛 | 2020/08/10 - 2020/09/24

MM Editing

# 整体结构



## 2.比赛方案介绍



# CenterNet for object detection

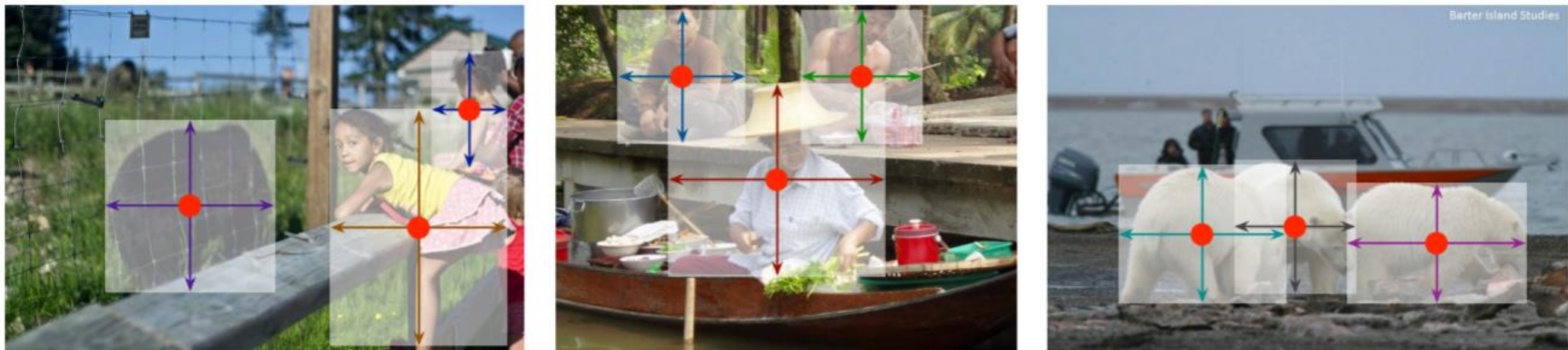
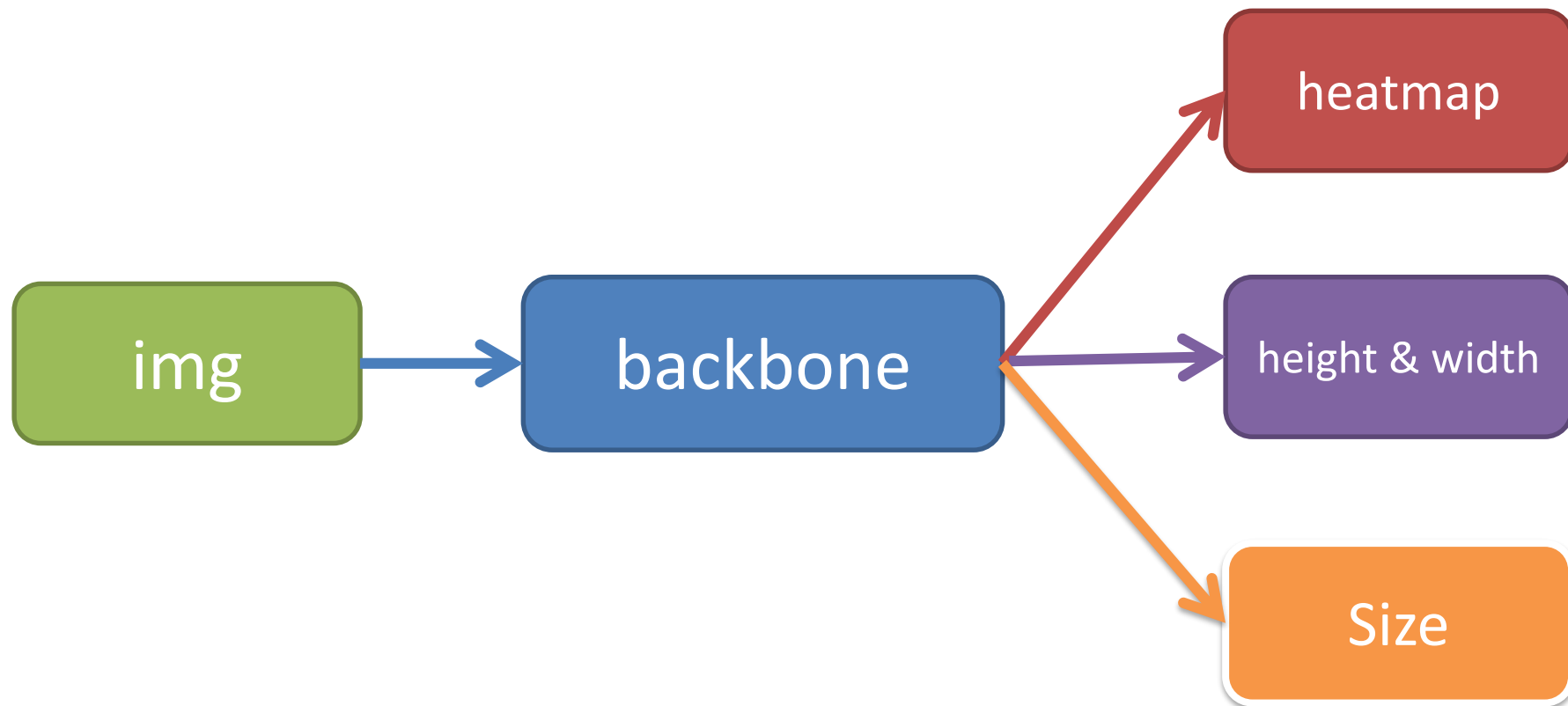
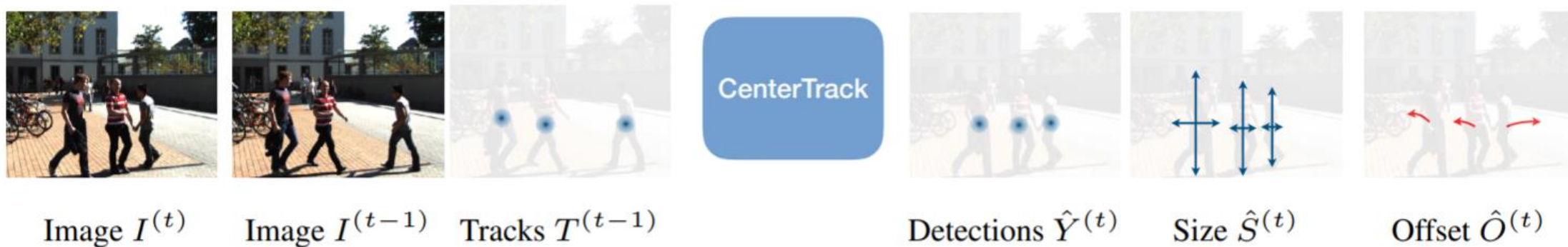


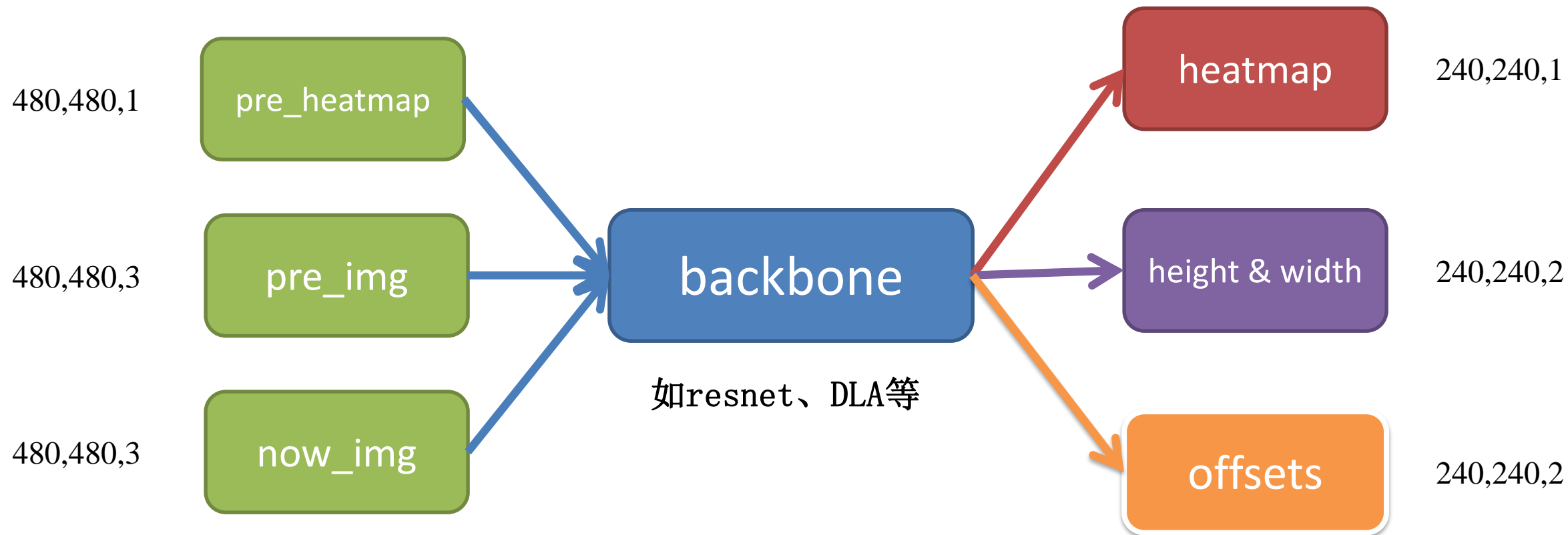
Figure 2: We model an object as the center point of its bounding box. The bounding box size and other object properties are inferred from the keypoint feature at the center. Best viewed in color.





# CenterTrack - Tracking Objects as Points





240,240,1

heatmap

240,240,2

height & width

240,240,2

offsets

$$\alpha = 2 \text{ and } \bar{\beta} = 4$$

$$L_k = \frac{1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}) & \text{if } Y_{xyc} = 1 \\ (1 - Y_{xyc})^\beta (\hat{Y}_{xyc})^\alpha \log(1 - \hat{Y}_{xyc}) & \text{otherwise} \end{cases},$$

$$L_{size} = \frac{1}{N} \sum_{i=1}^N |\hat{S}_{\mathbf{p}_i} - \mathbf{s}_i|.$$

$$L_{off} = \frac{1}{N} \sum_{i=1}^N \left| \hat{D}_{\mathbf{p}_i^{(t)}} - (\mathbf{p}_i^{(t-1)} - \mathbf{p}_i^{(t)}) \right|,$$

240,240,1

heatmap

240,240,2

height & width

240,240,2

offsets

## 定位:

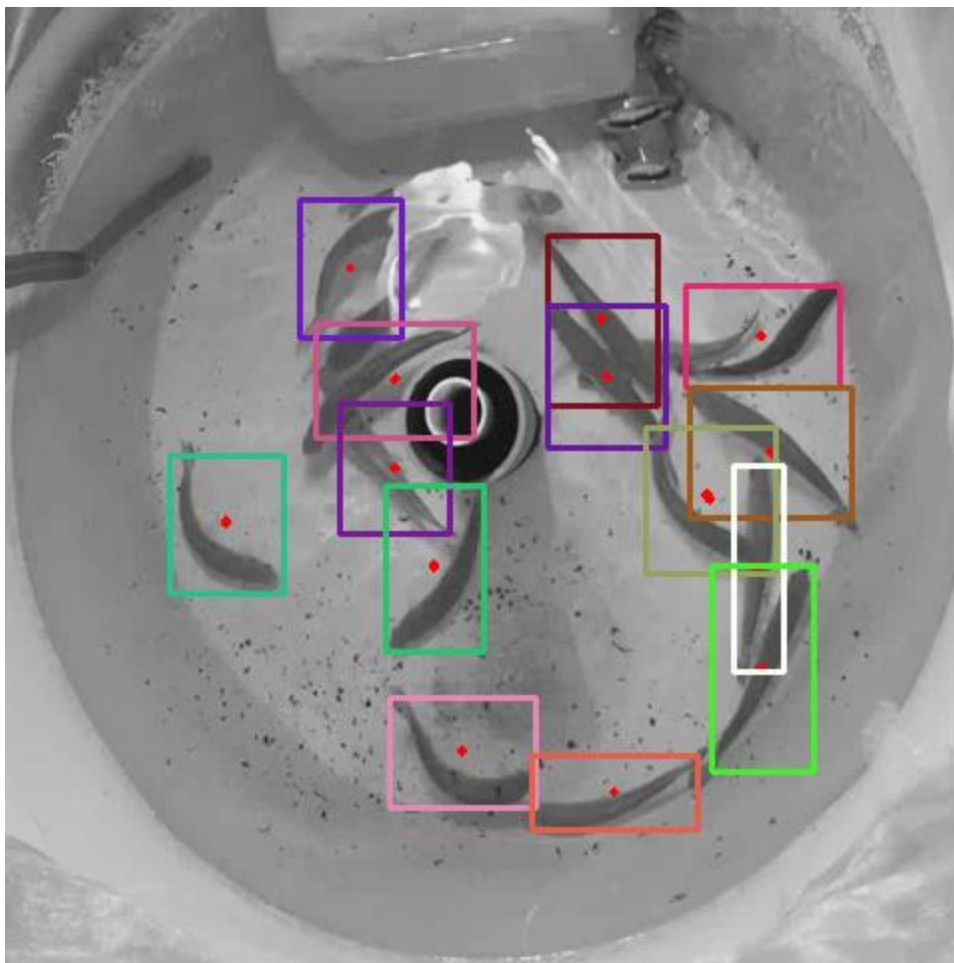
1. 根据heatmap, 设定阈值为0.27, 筛选出候选中心点
2. 使用5\*5 maxpooling进行非极大值抑制, 得到center坐标
3. 根据坐标去height&width中得到bbox的size

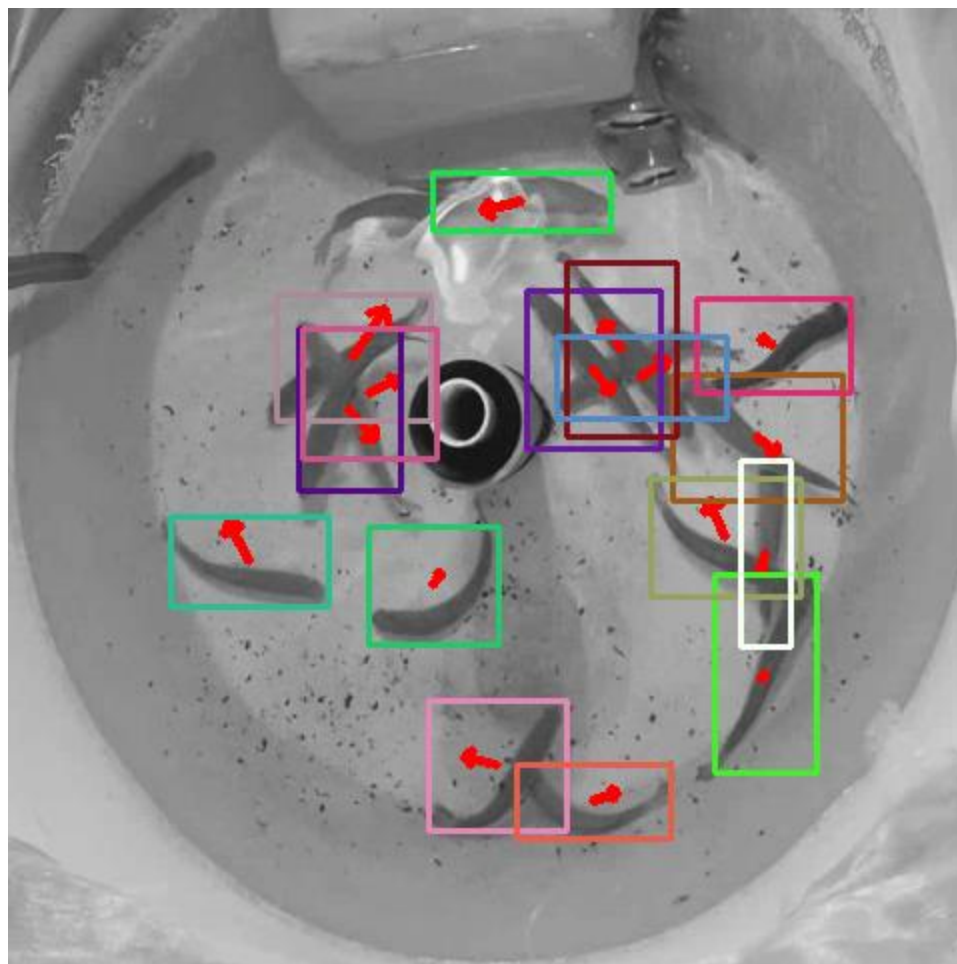
## 关联:

按照bbox置信度降序排序, 依次执行下列操作:

1. 根据center坐标去offsets中找到位移信息, 根据位移找到pre\_bboxes中与其距离最近的未关联bbox, 同时要求此距离要小于一定阈值T。
2. 若存在这样的pre bbox, 则为他们分配同样的id, 同时将pre bbox标记已经关联标记, 后续不再关联; 否则为当前bbox分配一个新id。

本比赛中,  $T=12$  (gap=1)     $T=20$  (gap=5)





- 使用DLA34(Deep layer aggregation)做为backbone，并在上采样层中使用DCNv2(Deformable Conv Network)，提高对运动的建模能力。
- backbone的stride为2，避免小鱼的漏检。
- 将backbone中的Batch Normalization替换为Group Normalization，多卡训练更稳定快捷。

<https://arxiv.org/abs/1803.08494>

<https://arxiv.org/abs/1707.06484>

<https://arxiv.org/abs/1811.11168>



- 对输入中的pre\_heatmap若使用GT，则会导致测试时存在domain gap，因此加入三种数据增强：
  - 一定概率在GT附近新加一个噪声bbox
  - 一定概率去除一个bbox
  - 对GT的bbox移动一小段距离（符合高斯分布）
- ColorJitter    brightness=0.4, contrast=0.6
- RandomCenterCropPad    ratios=(0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3)
- Flip
- pre gap in [1,2,3,4,5,6,7]      test gap = 1 or 5

## training:

5 \* TiTAN RTX                      80 epoch (8396 pair 1 epoch) with AdamW                      weightdecay =  $1e-4$   
batchsize 40                      learnin rate =  $5e-4$                       12小时

## test:

在TiTAN RTX上大约 8 FPS （没有针对性优化）， 模型完全在线（只需要前一帧信息）

## 展望:

- 使用过去更多帧的信息，理论上可以得到更稳定的跟踪结果
- 使用GIoU loss和更好的backbone，理论上可以得到更准确的检测结果

### 3. 一些代码细节

## 如何构造target heatmap (center loss label)

1. 对于每个object, 给一个iou阈值, 算出radius
2. 根据object的中心位置和radius修改heatmap
3. 修改heatmap时, 是取max不是覆盖

## 测试时，pre\_heatmap是怎么得到的

1. 拿上一帧检测的结果去构造（流程和训练时label构造一样，要保证相同的iou阈值）
2. 可以加一定的阈值（对结果再过滤一次）

## 4.未来展望

模型	状态	领域	发表
basicVSR	已复现相关指标	视频超分辨	cvpr2021
STTN	已复现模型，指标未对比	视频补全	eccv 2020
CenterTrack	已复现模型，指标未对比	多目标跟踪	eccv 2020
GIRAFFE	正在复现模型	3d & 生成	cvpr2021
.....	.....	.....	.....

# 感谢聆听

<https://github.com/Feynman1999/MgeEditing>