
Object localization using CNNs

1 人员分工

廖满文#: 实验思路的设计, 实验代码的书写

吴雨暄#: 实验代码的书写, 实验结果的可视化

饶 龙#: 实验代码的书写, 报告和 PPT 的制作

汪潇翔#: 报告和 PPT 的制作, 上台汇报

2 实验任务

数据集一共 5 类, 对于每一类用 150 张进行训练, 用 30 张作为测试集。

评估指标是计算要求在分类标签正确的情况下, IoU 需要大于 0.5 视为预测正确。

已完成内容:

在物体定位网络直接用 bbox 进行回归;

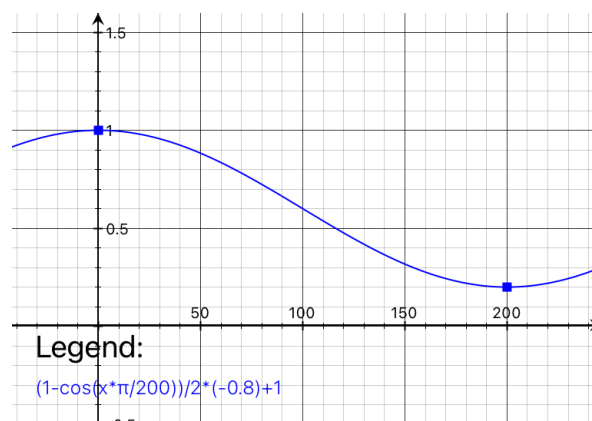
在网络中添加 anchors;

附加任务: 网络已经可以实现多目标定位了。

3 具体实现

本次模型的 Baseline 是使用了 ResNet50 作为 backbone;

学习率的设置是利用了 $1/4$ 个余弦函数, 初始学习率是 $1e-3$; 学习率的曲线如下所示;



优化器采用了 adam 优化器;



训练过程中的 batch_size 设置为 32;

损失函数分别设置为分类的损失函数为交叉损失熵, box 的损失函数设置为 MSE loss 函数, object 的损失函数设置为 BCEwithLogisticsLoss 函数;

训练过程中进行**数据增强**: 分别进行了平移, 旋转和尺度变化以及 cutout;

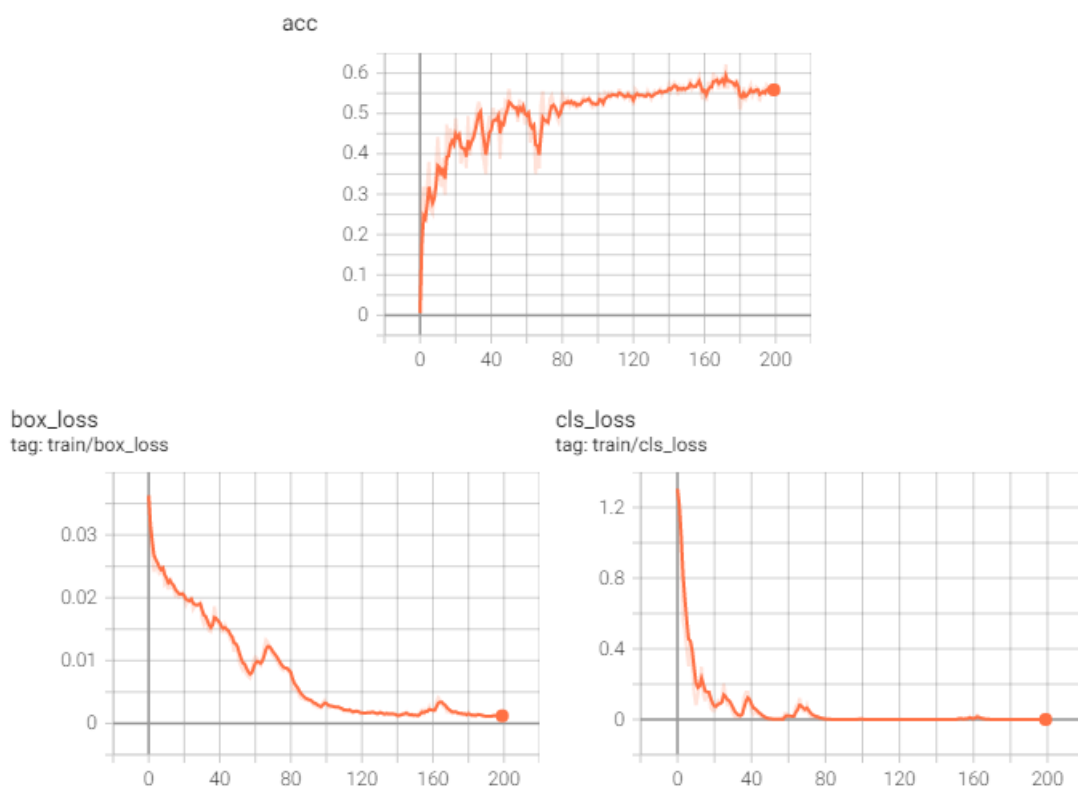
其中 anchors 的选取是将物体的 box 中心点进行 K-means 聚类聚成 9 类得到的;

使用了 grid cell 提高定位精度, 多目标定位可通过把 argmax 替换成 nms 实现。

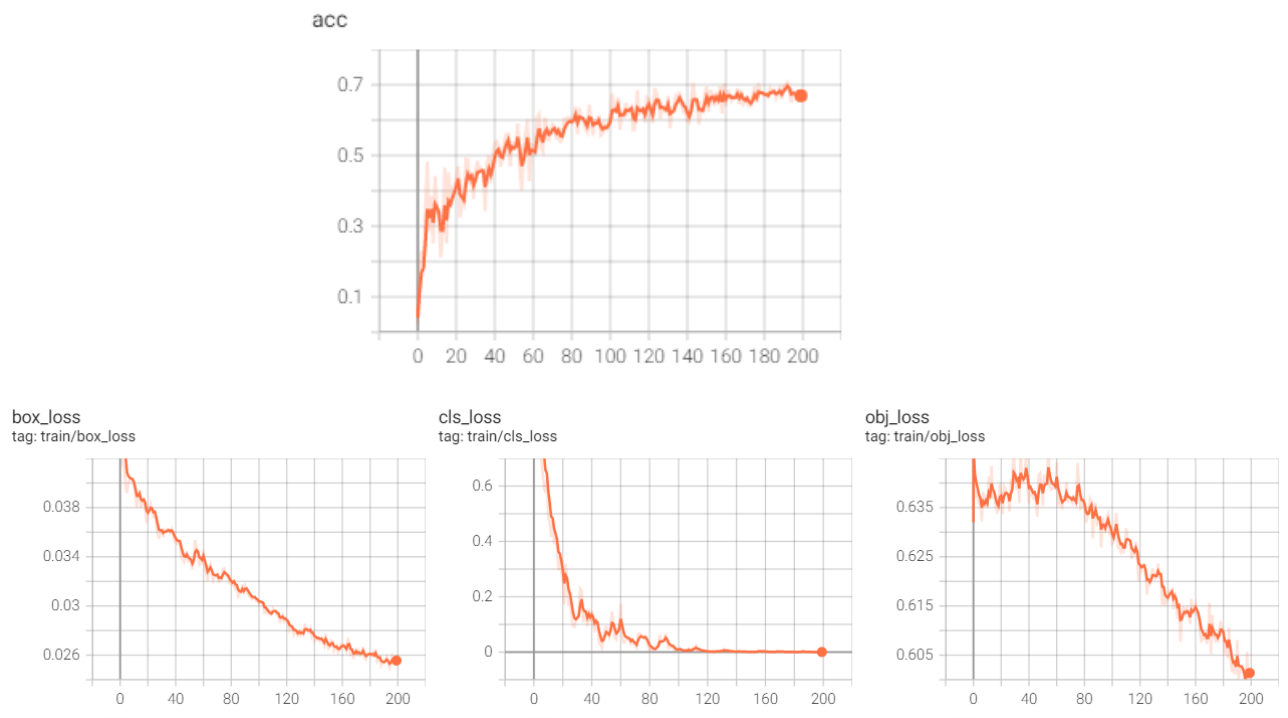
	ResNet-50	Cutout [3]
Image		
Label	Dog 1.0	Dog 1.0

4 实验结果

4.1 raw

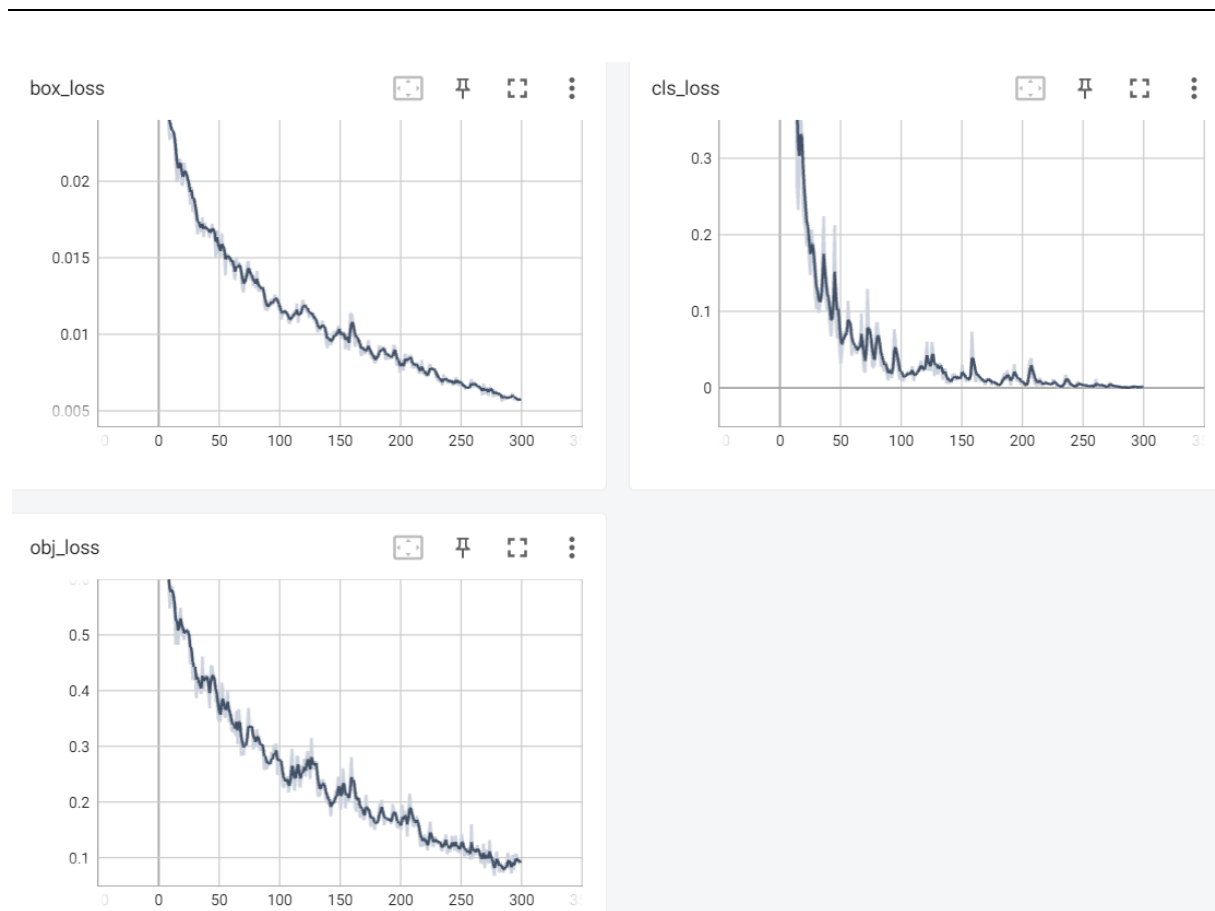


4.2 只加 data augment

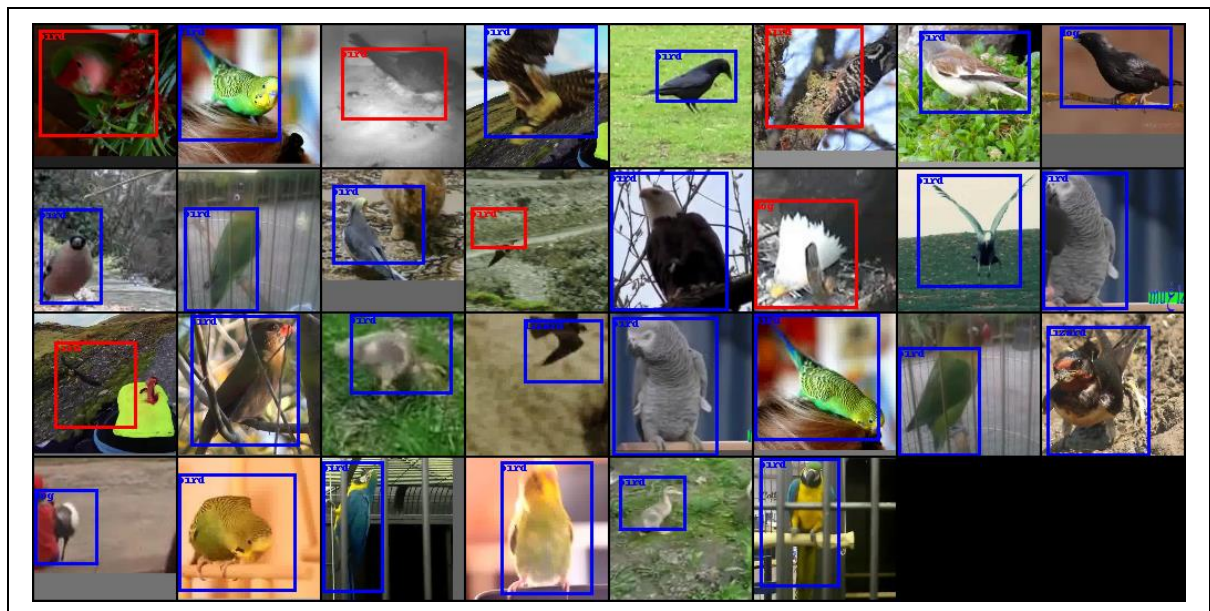


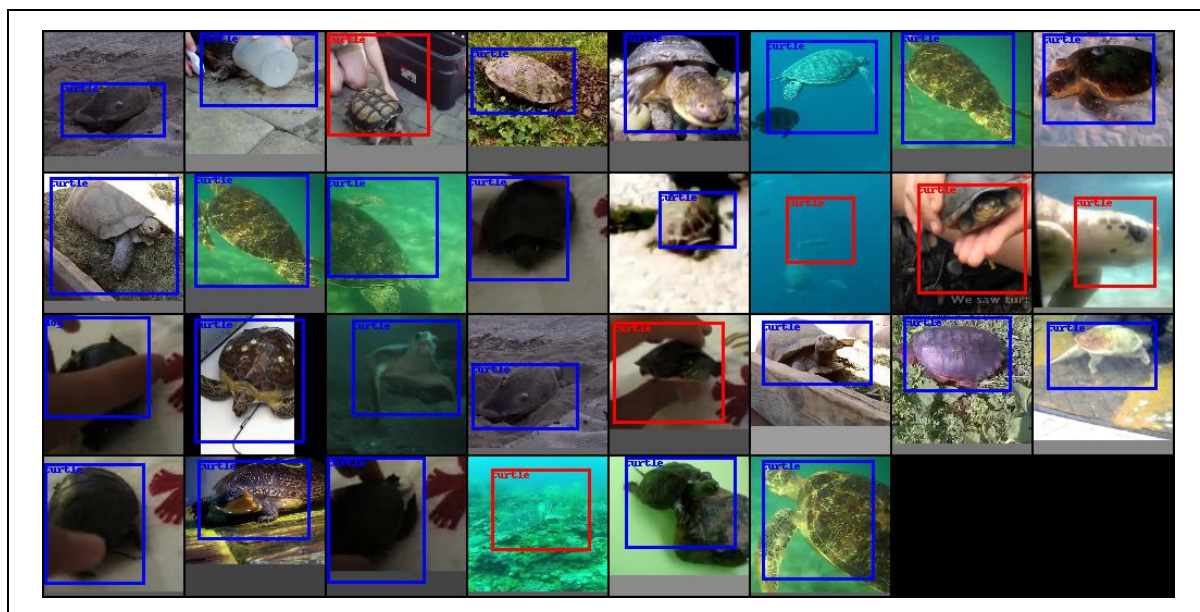
4.3 Anchor + data augment





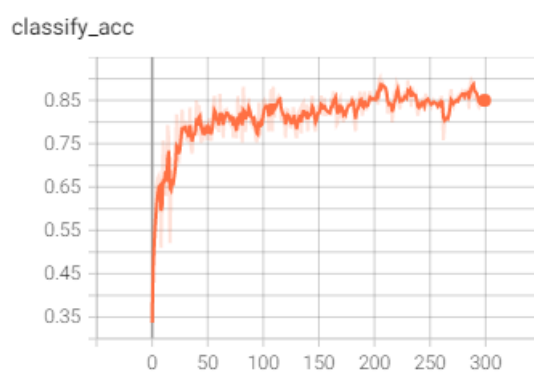
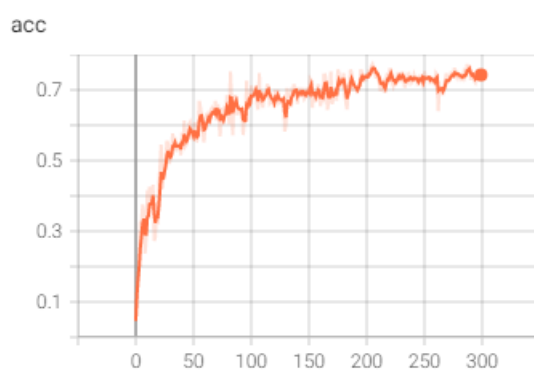
测试集的 30*5 张图片结果, 框上有预测出的类别, 图中黑框表示 $IOU > 0.5$, 红框表示 $IOU < 0.5$.

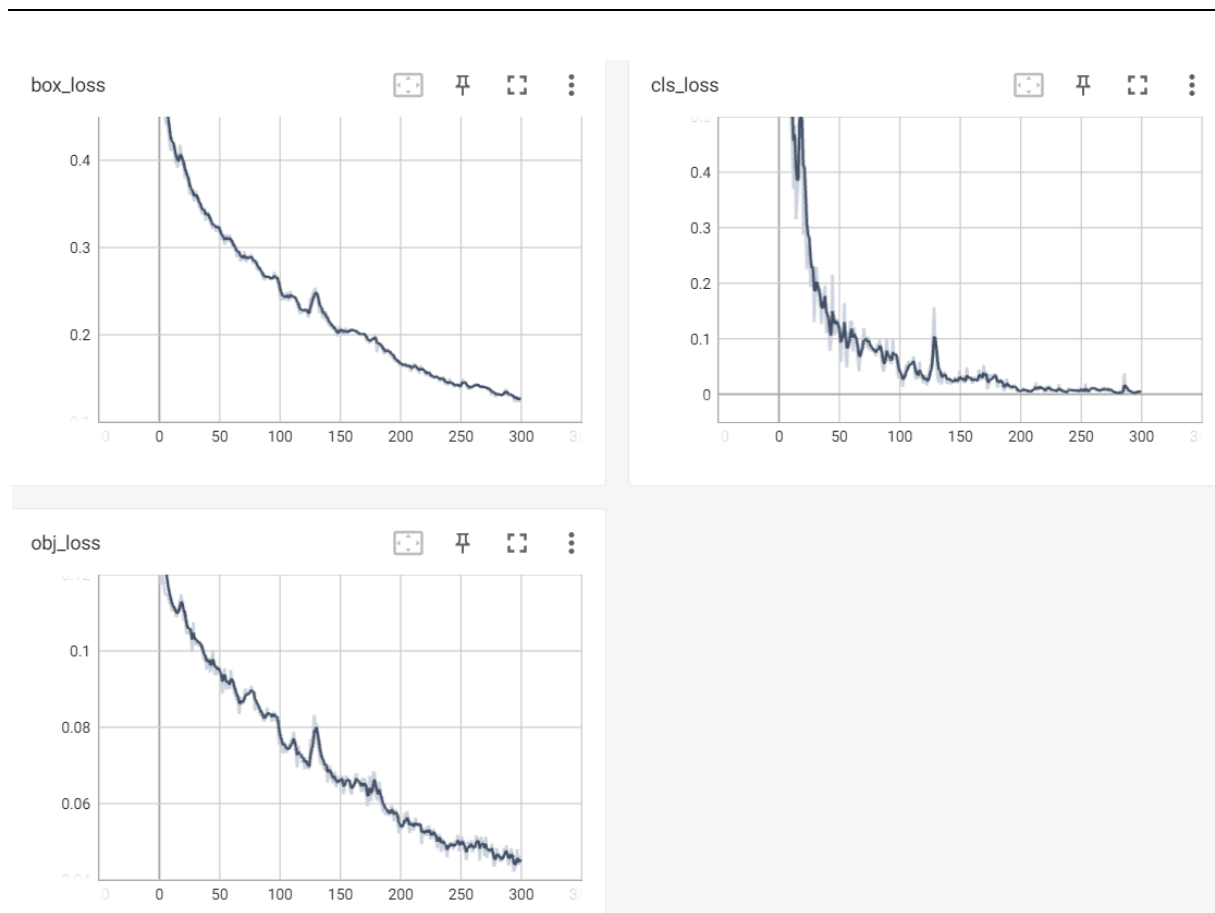




测试图像

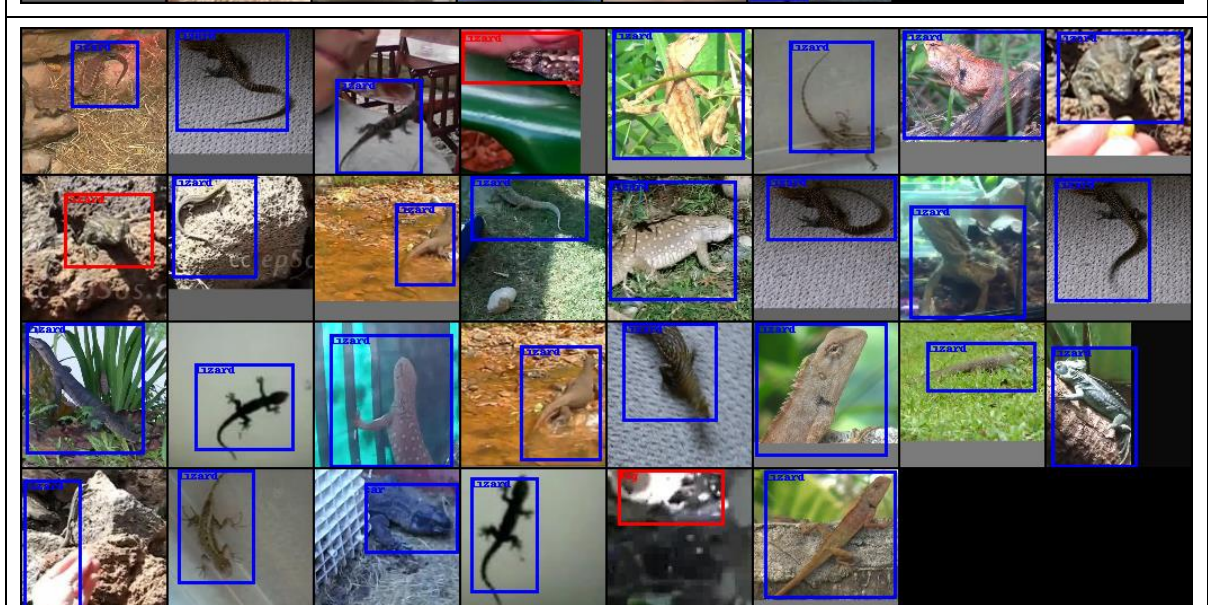
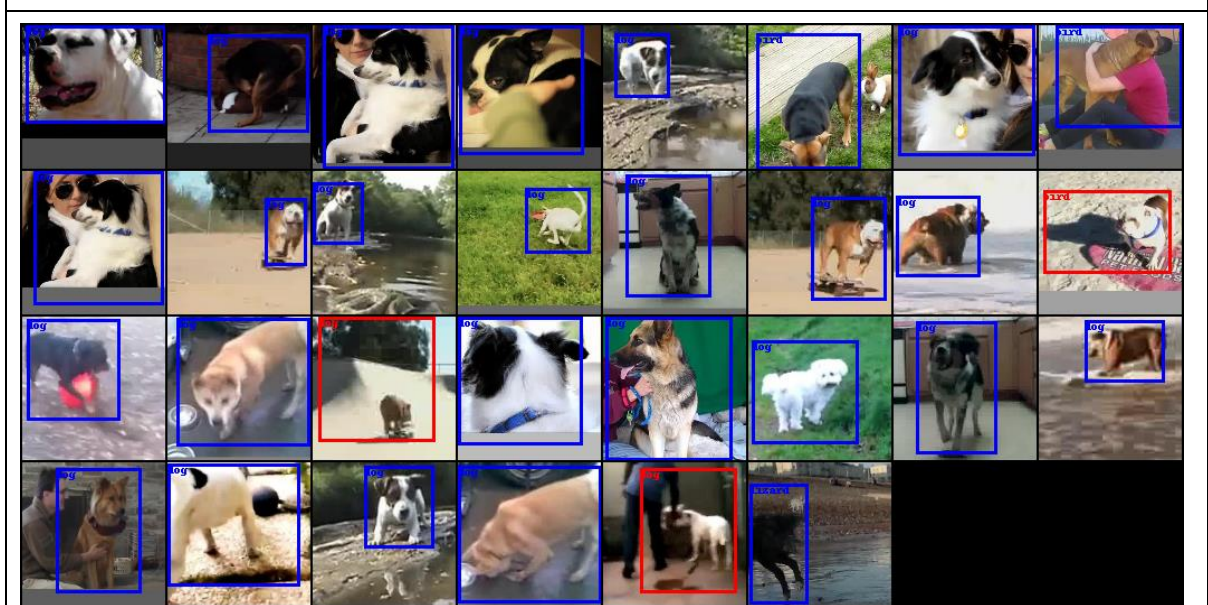
4.4 Anchor + make_grid + data augment

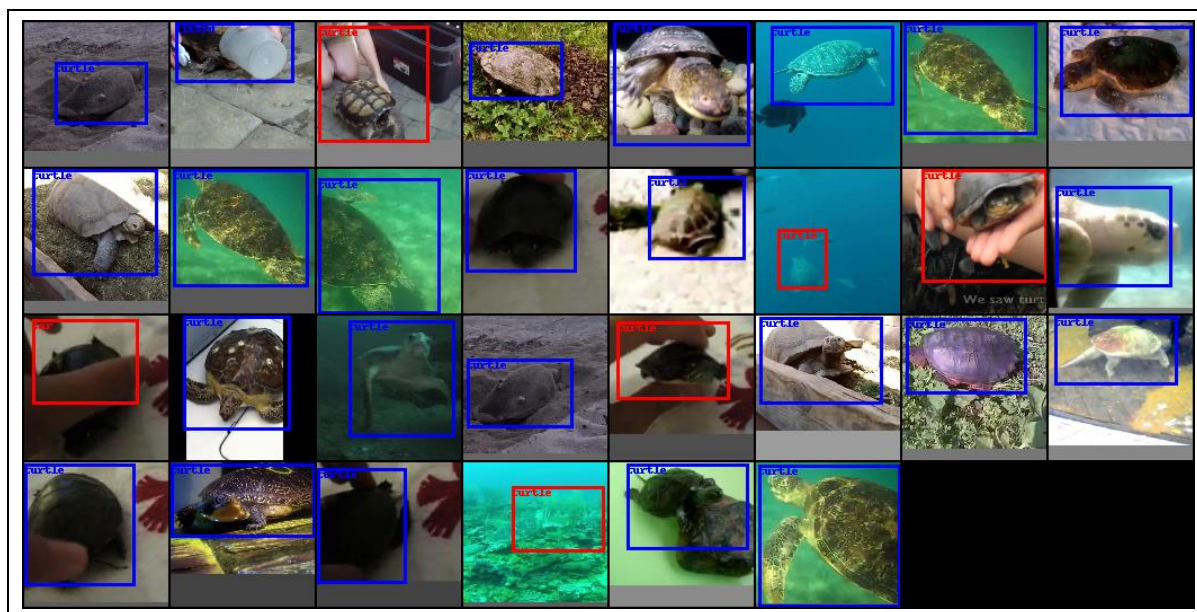




测试集的 30*5 张图片结果, 框上有预测出的类别, 图中黑框表示 $\text{IOU} > 0.5$, 红框表示 $\text{IOU} < 0.5$.







测试图像

5 总结

策略	Class_acc	Acc
raw		0.55
只加 data augment		0.67
Anchor + data augment	0.87	0.70
Anchor + make_grid + data augment	0.86	0.793

加入了 anchor 和 data augment 以及 make_grid 之后，实验的精度会明显提高。最终在测试集上的，实验结果将达到 0.793