

Gaussian YOLOv3: An Accurate and Fast Object Detector Using Localization Uncertainty for Autonomous Driving

笔记本： 文件中转

创建时间： 2019/11/3 19:12

更新时间： 2019/11/5 19:56

作者： ming71

ICCV 2019

解决的问题比较经典，方法比较有趣，但是效果其实有限而且不具有很强的insight，看看就行。

1. Introduction

- 研究背景

自动驾驶中的车辆检测问题

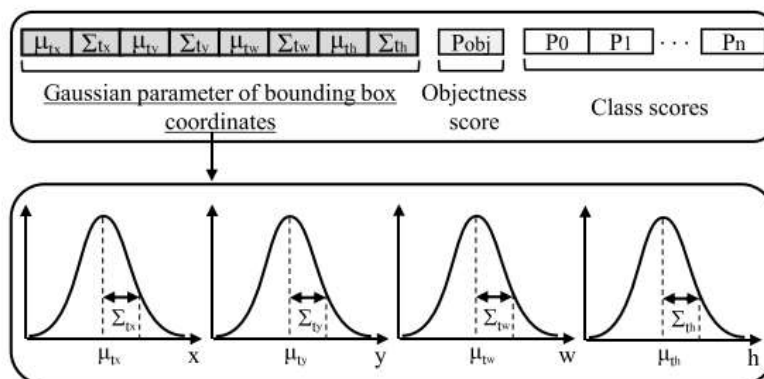
- 提出问题

研究的问题仍是IoUNet指出的，用分类分数指导bbox的NMS不科学的问题，这里换了种表述，说是没有评价bbox预测的不确定性（即默认所有预测出来的bbox置信度为1）。作者着眼的地方是FP（false positive）误检样本，归因于回归的时候没有考虑bbox的不确定性（本质上是以分类作为box的不确定性来指导选择的）

- 提出方案

预测的bbox不再是简单直接的坐标，而是四个预测值（xywh）的高斯分布，引入对bbox的不确定性评估，随之loss也要改。

2. Gaussian YOLOv3



思路比较直观，直接将原来的xywh预测改为四个值的每个值预测一对高斯参数。将输出的参数xy进行sigmoid归一化，即可得到高斯分布的均值和方差，其中均值和预测bbox的真值相关联（一个grid cell内的偏移），方差关联bbox的不确定性（未经过sigmoid）。

关于loss: 对于分类和obj置信损失未更改，只是调整了bbox定位的loss。对应下面公式，看上去比较复杂，核心思想就是对于给定的高斯分布，求取对应位置的概率作为不确定性来衡量相应回归信息的loss。

$$L_x = - \sum_{i=1}^W \sum_{j=1}^H \sum_{k=1}^K \gamma_{ijk} \log(N(x_{ijk}^G | \mu_{t_x}(x_{ijk}), \Sigma_{t_x}(x_{ijk})) + \epsilon),$$

定位不确定性：相比于yolov3的分类分数乘以置信度作为分类不确定性的衡量，这里再乘了个高斯分布取值的不确定度。

$$Cr. = \sigma(Object) \times \sigma(Class_i) \times (1 - Uncertainty_{aver}).$$

3. Experiment

Detection algorithm	Average precision (%)									mAP (%)	FPS	Input size
	E	Car										
		M	H	E	M	H	E	M	H			
MS-CNN [1]	92.54	90.49	79.23	87.46	81.34	72.49	90.13	87.59	81.11	84.71	8.13	1920×576
SINet [11]	99.11	90.59	79.77	88.09	79.22	70.30	94.41	86.61	80.68	85.42	23.98	1920×576
SSD [17]	88.37	87.84	79.15	50.33	48.87	44.97	48.00	52.51	51.52	61.29	28.93	512×512
RefineDet [28]	98.96	90.44	88.82	84.40	77.44	73.52	86.33	80.22	79.15	84.36	27.81	512×512
CFENet [29]	90.33	90.22	84.85	-	-	-	-	-	-	-	0.25	-
RFBNet [16]	87.41	88.35	83.41	65.85	61.30	57.71	74.46	72.73	69.75	73.44	39.20	512×512
YOLOv3 [21]	85.68	76.89	75.89	83.51	78.37	75.16	88.94	80.64	79.62	80.52	43.57	512×512
Gaussian YOLOv3	90.61	90.20	81.19	87.84	79.57	72.30	89.31	81.30	80.20	83.61	43.13	512×512
Gaussian YOLOv3	98.74	90.48	89.47	87.85	79.96	76.81	90.08	86.59	81.09	86.79	24.91	704×704

实验来看效果还可以，但是没有特别的惊艳。从原理来看，想要解决的问题是bbox的不确定性，这在单阶段检测上有一定的意义（但是这种mismatch感觉没有那么严重，往往分类好的检测效果都还行，就算IoUNet也是用了cascade回归和lou预测作为创新点），但是效果有限；

实现的方式很有趣，是学习一个高斯分布来拟合这个不确定性的分布

