

# 专业阅读与写作第四次作业

58119304 朱启鹏

## I. 感兴趣的主题

a) 主题：：目标检测

b) 理由：：(1) 自从神经网络掀起人工智能的一场春风，CV 领域发展的势头越来越汹涌，是当今人工智能领域一大热门方向。能投身于计算机视觉领域，进行相关的科研工作，将使我不胜荣幸。其中，目标检测是 CV 的一块重要的子领域，且在自动驾驶、医学图像、物体检测都具有广泛的应用价值

(2) 另一方面，计算机视觉的工作虽然繁杂，且具有多交叉、综合性强的特点，但是其成果具有极强的可视化，可以很好的展示给工作者的眼前，这能够很好的激励工作者的工作。简言之，这是一项极具挑战性的工作，但是与之带来的是很强的正反馈，这压力与动力并存的任务，能够使我更好地砥砺前行。

(3) 最后，我最终地方向希望朝着视频有关方向地工作。所以，先选择目标检测进行入手，将是一个不错的选择。

## II. 最新高水平论文

a) 题目：[4]DAP: Detection-Aware Pre-training with Weak Supervision

b) 简介：(1) 本文提出了一种检测感知预训练方法，该方法仅利用弱标记的分类样式数据集进行预训练，但专门为使目标检测任务受益而量身定制。与广泛使用的基于图像分类的预训练不同，它不包括任何与位置相关的训练任务，本文通过基于类激活图的弱监督对象定位方法将分类数据集转换为检测数据集，直接预训练检测器，使预先训练的模型具有位置感知能力并能够预测边界框。在下游检测任务中，DAP 在效率和收敛速度方面都可以优于传统的分类预训练。特别是当下游任务中的样本数量很少时，DAP 可以大幅提高检测精度。

(2) 实验步骤展示：一般来说，检测器的主干网络都需要在 ImageNet 之类的分类数据集上预训练。同时分类标签也可以用来学习目标检测，这个情况下称之为弱监督目标检测。如果把预训练的图像分类任务难度提高成目标检测，就有了这篇文章，原理一图就能看明白。

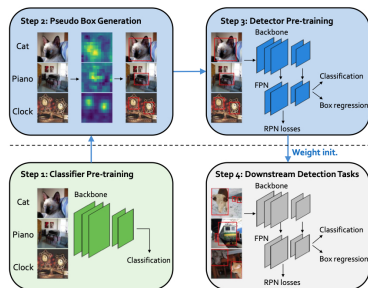


Fig. 1. 步骤

(3) 工作流程预览：

1. 模型与训练

2. 用 CAM 生成 Pseudo Box

$$\text{CAM}(x, y) = \frac{\text{CAM}(x, y) - \min \text{CAM}(x, y)}{\max \text{CAM}(x, y) - \min \text{CAM}(x, y)} \quad (1)$$

$$M(x, y) = \text{CAM}(x, y) \times \mathbb{I}\{\text{CAM}(x, y) > \tau\} \quad (2)$$

$$x_c = \frac{\sum_{(x,y) \in \Omega} M(x, y)x}{\sum_{(x,y) \in \Omega} M(x, y)} \quad (3)$$

$$y_c = \frac{\sum_{(x,y) \in \Omega} M(x, y)y}{\sum_{(x,y) \in \Omega} M(x, y)} \quad (4)$$

$$w = \sqrt{12 \frac{\sum_{(x,y) \in \Omega} M(x, y) (x - x_c)^2}{\sum_{(x,y) \in \Omega} M(x, y)}} \quad (5)$$

$$h = \sqrt{12 \frac{\sum_{(x,y) \in \Omega} M(x, y) (y - y_c)^2}{\sum_{(x,y) \in \Omega} M(x, y)}} \quad (6)$$

3. 检测器预训练

4. 下游探测器微调

(4) 实验结果：

TABLE I  
COCO FULL-DATA 检测结果

Pre-train	AP <sub>5:95</sub>	AP <sub>.5</sub>	AP <sub>.75</sub>	AP <sub>s</sub>	AP <sub>m</sub>	AP <sub>l</sub>
IN-1M CLS	36.73	58.04	39.72	20.57	39.56	48.51
IN-1M DAP	37.25	58.98	40.46	21.71	40.64	48.34
$\Delta$	+0.52	+0.94	+0.74	+1.14	+1.08	+0.83
IN-14M CLS	38.87	61.87	42.41	23.79	42.15	49.89
IN-14M DAP	39.57	63.05	43.02	24.03	42.96	51.15
$\Delta$	+0.70	+1.18	+0.61	+0.24	+0.81	+1.26

## III. 关键词，重要论文，以及学者

a) 关键词：深度卷积网络，目标检测，有/弱/自监督，点云

理由：目前采取目标检测的技术主要为深度卷积网络，而对于数据的有无标签，其研究可以分为有监督与无监督学习，由此衍生而来的便是弱监督和自监督

b) 重要论文：[1]A hierarchical graph network for 3D object detection on point clouds

[2]Point-gnn: Graph neural network for 3d object detection in a point cloud

[3]Hvnet: Hybrid voxel network for lidar based 3d object detection

理由：这几篇文章所做的工作为后续研究提供了良好的思路与方法，并且是现在最前沿的论文，十分具有研究意义。更重要的是，这三篇论文均为 2020 年 CVPR 的作品，非常具有研究价值。

c) 著名学者: 何恺明、孙剑、任少卿、Xiangyu Zhang、李飞飞、汤晓鸥

理由: 为了能够真实地反映人工智能领域学者的研究水平以及国家发展水平, 1 月 11 日, 清华-中国工程院知识智能联合研究中心和清华大学人工智能研究院推出了「AI 2000 人工智能全球 2000 位最具影响力学者榜单 (AI 2000 Most Influential Scholar Award)」。

排在计算机视觉领域第一名的是著名青年学者何恺明, 入选论文总引用量 70524。随之其后的是孙剑、任少卿、Xiangyu Zhang, 入选论文总引用量分别为 64419、50104、47322。值得一提的是计算机视觉领域的知名学者李飞飞和汤晓鸥分别以入选论文总引用量 31856 和 27521 分别排名第 8 名和第 10 名。

#### REFERENCES

- [1] Jintai Chen, Biwen Lei, Qingyu Song, Haochao Ying, Danny Z Chen, and Jian Wu. A hierarchical graph network for 3d object detection on point clouds. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 392–401, 2020.
- [2] Weijing Shi and Raj Rajkumar. Point-gnn: Graph neural network for 3d object detection in a point cloud. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1711–1719, 2020.
- [3] Maosheng Ye, Shuangjie Xu, and Tongyi Cao. Hynet: Hybrid voxel network for lidar based 3d object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1631–1640, 2020.
- [4] Yuanyi Zhong, Jianfeng Wang, Lijuan Wang, Jian Peng, Yu-Xiong Wang, and Lei Zhang. Dap: Detection-aware pre-training with weak supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4537–4546, 2021.