**基于深度学习的小型无人机视觉检测技术研究**

**1.题目背景：**

近年来随着无人机技术的飞速发展，无人机被广泛应用于多种领域，如航拍、运输、救援等。同时，无人机的黑飞、滥飞现象也层出不穷，如擅闯机场禁飞区、拍摄居民楼内隐私照片等，对公共安全和个人隐私造成极大的危害。因此，通过视觉检测技术监测敏感空域中的无人机迫在眉睫。常见的民用无人机体型较小，而无人机监测系统所需求的探测距离较大，因此无人机在图像画面中的像素面积较小，对无人机视觉监测系统的小目标检测能力提出严峻挑战。

**2.题目介绍：**

本题目希望帮助同学们了解经典的卷积神经网络，并对无人机的视觉检测技术展开研究，并进一步引导同学们针对如何提升无人机小目标的检测能力展开探索。**题目分为两个部分：1.无人机检测基础任务；2.无人机小目标检测进阶任务。**

**2.1 无人机检测基础任务（必做）：**

常用于无人机视觉检测的经典卷积神经网络有Faster R-CNN [1], YOLOv3 [2], CenterNet [3] 等等（均有开源代码）。同学们需自选一种经典的卷积神经网络并阅读相关论文以加深了解，然后基于给定的无人机检测数据集，将对应网络模型部署到无人机检测任务上。具体而言，题目给定包含小型无人机的数据集，且数据集被划分为两个部分：训练集和测试集。同学们需在训练集上训练自己的深度学习模型，并在测试集上测试其检测精度(AP at IoU=0.50) [4]，**不可以使测试集参与训练**。完成无人机数据集上的仿真验证实验后，同学们需在提交文档中指出自己所选用的经典卷积神经网络模型中，作者设计了哪些结构或者采用了哪些Tricks有助于无人机小目标检测，指出一两点即可。

**2.2 无人机小目标检测进阶任务（选做）：**

本题中的**小目标定义**为：**无人机轮廓框的像素面积小于32\*32像素** [4]。

同学们需自我探索如何在自己所选用的经典卷积神经网络基础上提升对无人机小目标的检测效果，可以考虑的方法包括但不限于：

1. 数据增强：通过随机剪切粘贴等方式扩增数据集中的小目标数量；
2. 迁移训练：先在中等以上大小的目标上训练模型，然后再在小目标上进行针对训练；
3. 多尺度特征融合（常用）：通过特征金字塔、路径聚合等方式将小感受野的特征图和大感受野的特征图进行融合，同时提升小目标和大目标的检测能力；
4. 调参（简单、直接）：提高Region Proposals [1]、Anchors [2]、Centers[3]的覆盖密度；
5. 超分辨率特征增强（困难）：设计超分辨率特征增强模块，使网络对小目标特征更为敏感。

**3.题目细节：**

**3.1 数据集**

1. 考虑到计算资源受限，我们选择了南加州大学（USC）的旋翼无人机数据集包含1765张图片和对应的1765个txt标签。USC数据集不能商用，我们希望同学们不要向外传播。

链接： https://pan.zju.edu.cn/share/4b3d6ec2281c6167a39deb2e7d

访问密码：1234

2. 标签为文件夹内的与图片同名的txt格式，一行包含五个数字：l x y w h。l代表类别，在本数据集中都为0，即代表无人机。x y w h用来描述无人机轮廓框在图片中的位置，并且都以图片长宽的比例形式表征。假设图片长度为W，宽度为H，以图片的左上角为原点，向右为x正方向，向下为y正方向，则轮廓框的中心坐标为（x\*W，y\*H），轮廓框的长度为w\*W，轮廓框的宽度为h\*H。

3. 请同学们自行按照4:1的比例将数据集随机划分为训练集和测试集，用训练集训练模型，用测试集评估模型的检测精度（AP at IoU=0.50，AP指标请参考MSCOCO数据集[4]），注意不能让测试集参与训练过程。

**3.2 最终提交内容**

最终提交一份Word文档，需包含以下内容：

1. 必做部分：对自己所选用卷积神经网络模型的简要介绍，不要原文翻译；
2. 必做部分：模型训练和测试的实验记录；
3. 必做部分：简要说明所选用经典卷积神经网络的作者设计了哪些结构或者采用了哪些Tricks有助于无人机小目标检测；
4. 选做部分：详细描述自己如何在经典网络基础上增强其对无人机小目标的检测能力，并附上消融实验记录（原网络和改进后网络对于小目标的检测精度对比）

**3.3 评分规则**

最终提交文档的必做部分（1~3）构成基础分，必做部分不看重模型的检测精度，同学们不必刻意刷分；而最终提交文档的选做部分（4）作为加分项。

**4.参考资料**

[1]Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28: 91-99.

[2] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[3] Zhou X, Wang D, Krähenbühl P. Objects as points[J]. arXiv preprint arXiv:1904.07850, 2019.

[4] Lin T Y, Maire M, Belongie S, et al. Microsoft coco: Common objects in context[C]//European conference on computer vision. Springer, Cham, 2014: 740-755. Details at: https://cocodataset.org/