

CV项目实践赛

以下是一些建议的可实践项目，当然欢迎大家提出更多新的创意，只要项目方向与计算机视觉相关且具有实际应用意义。可以使用公开数据集，也鼓励同学们使用自建数据集。模型方面，除了可以使用现有模型或预训练模型外，也欢迎进行算法创新，但需确保符合常规的训练或微调流程。

项目总览

项目	领域	核心任务	数据规模	模型参考
肺炎X光分类	医疗	二分类	~5,863	MobileNetV3, EfficientNet-B0
街景道路分割	自动驾驶	语义分割	~500	U-Net+ResNet18, DeepLabv3+
植物病害检测	农业	多分类	~2,000	YOLOv5s, ViT-Tiny
卫星建筑分割	遥感	语义分割	360	Fast-SCNN, FPN
零售货架检测	商业	目标检测	~2,000	YOLOv8n, SSD-MobileNet
视网膜血管分割	医疗	语义分割	40	U-Net+Attention, LinkNet
番茄果实实例分割	农业	实例分割	2,842	Mask R-CNN, YOLOv8-seg

项目1：肺炎X光图像分类

背景：辅助医生快速筛查胸部X光片中的肺炎症状

数据集：

- ChestX-ray8 (5863张X光片，2分类：正常/肺炎)
- 来源：<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

参考模型：

- MobileNetV3（轻量级分类，冻结预训练层后微调）

- EfficientNet-B0（平衡精度与速度）

难度：★★☆（需处理医学影像不平衡问题）

小贴士：医学图像通常存在类别不均衡的情况，特别是肺炎X光图像。医学图像通常存在类别不均衡问题。建议使用重复采样（过采样阳性样本），结合弹性形变、随机旋转等医学图像增强技术，并采用加权交叉熵损失（如阳性样本权重设为3-5倍）来缓解类别不平衡问题，确保模型能够在阳性样本较少的情况下依然保持高准确性。

项目2：街景道路语义分割

背景：自动驾驶中的道路可行驶区域识别

数据集：

- Cityscapes子集（选取500张精细标注的城市街景图，19类语义标签）
- 来源：<https://www.cityscapes-dataset.com/>

参考模型：

- U-Net with ResNet18 backbone（适合小样本分割）
- DeepLabv3+ Mobile（轻量级实时分割）

难度：★★★（需处理多类别像素级标注）

小贴士：Cityscapes数据集包含19类语义标签，处理时需要特别注意多类别分割问题。在选择500张子集时，要确保每个类别的标注都有足够的样本，以避免类别不平衡影响模型表现。可以使用加权损失函数来帮助模型更好地学习少数类。

项目3：植物叶片病害分类

背景：农业自动化中的作物健康监测

数据集：

- PlantVillage Dataset（5.4万张叶片图像，38类病害/健康状态）
- 可选取番茄、马铃薯等3-5类子集（约2000张）
- 来源：<https://plantvillage.psu.edu/>

参考模型：

- Vision Transformer Tiny（小规模图像分类）
- EfficientNet-B3（平衡精度与速度）

难度：★★☆（需平衡不同病害样本量）

小贴士：在PlantVillage数据集中，病害的类别不均衡可能会影响模型的学习效果。可以尝试使用数据增强（如旋转、裁剪、颜色扰动等）以及损失函数加权，特别是对少数类进行过采样或使用合成数据技术（如SMOTE），以增强模型的泛化能力。

可参考：https://github.com/MilleXi/plant_diseases_recognition

项目4：卫星图像建筑物分割

背景：城市规划中的建筑物分布分析

数据集：

- Inria Aerial Image Labeling Dataset（360张卫星图，512x512分辨率）
- 来源：<https://project.inria.fr/aerialimagelabeling/>

参考模型：

- FPN (Feature Pyramid Network)
- Fast-SCNN（实时轻量分割网络）

难度：★★★（需处理高分辨率图像分块训练）

小贴士：由于Inria Aerial数据集包含的图像较少（360张），分割任务可能存在过拟合风险。可以考虑使用数据增强技术（如翻转、旋转、裁剪等）来增加数据多样性，或考虑选择更大规模的子集进行训练，以提高模型的鲁棒性。

项目5：零售货架目标检测

背景：零售商品自动盘点与陈列分析

数据集：

- SKU-110k（11,762张货架图片，标注商品边界框）
- 可选取含饮料/零食的2000张子集
- 来源：https://github.com/eg4000/SKU110K_CVPR19

参考模型：

- SSD with MobileNet backbone
- YOLOv8n（Nano版适合低算力）

难度：★★☆（需处理密集小目标检测）

小贴士：SKU-110k数据集适合目标检测任务，但由于图像中的商品密集且目标小，训练时可能需要专注于密集小目标的检测。训练时建议将输入分辨率调整为**640x640**，

并启用**多尺度训练**（如0.5x-1.5x缩放），也可以通过适当选择Anchor Box的尺寸、使用Focal Loss（针对小目标）和改进NMS（非极大值抑制）来提高检测精度。

项目6：视网膜血管医学分割

背景：眼科疾病诊断（如糖尿病视网膜病变）中的血管结构分析

数据集：

- DRIVE Dataset（40张视网膜图像 + 血管标注mask）
- 特点：专业医学标注，包含测试集/训练集划分
- 来源：<https://drive.grand-challenge.org/>

参考模型：

- U-Net with Attention Gate（提升血管细节捕捉能力）
- LinkNet（轻量级实时分割架构）

难度：★★★☆☆（医学图像对比度低，需精细结构调整）

小贴士：DRIVE数据集的图像通常对比度较低，血管结构不易分辨。为了提高模型的分割精度，可以采用图像增强技术，如直方图均衡、CLAHE（自适应直方图均衡）等，以提升图像对比度和细节表现，从而帮助模型更好地识别细小的血管结构。

可参考：<https://github.com/MilleXi/Retinal-vessel-segmentation>

项目7：番茄果实实例分割

背景：农业自动化中的果实成熟度检测与产量预估

数据集：

- Laboro Tomato Dataset（2,842张番茄图像，实例分割标注）
- 标注类型：边界框 + 像素级mask（健康/病变/成熟度分级）
- 来源：<https://github.com/laboroai/Laboro-Tomato>

参考模型：

- Mask R-CNN with ResNet50-FPN（经典实例分割框架）
- YOLOv8-seg（轻量级实时分割，Nano版本）

难度：★★★（需同时处理检测与分割任务）

小贴士： Laboro Tomato数据集中，果实可能会出现重叠现象，导致检测困难。为了应对这种情况，可以优化NMS（非极大值抑制）算法，选择更合适的IoU阈值来避免误检，同时进行数据增强（如平移、缩放）来增强模型对果实重叠情况的处理能力。

关键实施建议（面向所有项目）

1. 数据优化

- **增强策略：**
 - 医学图像：CLAHE对比度增强 + 弹性形变（DRIVE项目）
 - 农业/街景：Albumentations组合增强（旋转/裁剪/色彩抖动）
- **样本均衡：**
 - 过采样少数类（如肺炎X光中的阳性样本）
- **分块训练：** 高分辨率图像（卫星/视网膜）切分为256x256子图

2. 模型选择

- **小数据集：** 优先使用预训练模型（ImageNet权重迁移）
- **医学分割：** 选择U-Net变体（Attention U-Net/ResUNet）
- **实时需求：** 轻量模型（YOLOv8n、DeepLabv3+ Mobile）

3. 训练优化

- **基础配置：**
 - 学习率：`1e-4` ~ `3e-5`（分割任务更低）
 - 损失函数：Dice Loss（分割）、Focal Loss（检测类别不均衡）
- **加速收敛：**
 - 启用早停法（`patience=5`） + 混合精度训练
 - 梯度累积（模拟大batch_size）

4. 资源管理

- **显存不足：**
 - 降低 `batch_size` 至4-8
 - 启用梯度检查点（`torch.utils.checkpoint`）

- **免费算力**：Colab/Kaggle GPU + 开启 `-batch=16 --workers=2`
-

执行优先级建议

1. **入门推荐**：肺炎X光分类（任务简单，数据量大） → 零售货架检测（密集目标实践）
2. **进阶挑战**：视网膜血管分割（医学精度要求高） → 番茄实例分割（检测+分割联合任务）
3. **避坑提示**：卫星图像需分块训练避免OOM；Laboro Tomato需软化NMS处理果实重叠

所有项目代码均可通过PyTorch/Keras在 **单卡GPU（4GB+显存）** 完成，平均训练时间约2-4小时/项目阶段。