

人工智能实验 团队大作业

苹果香蕉大作战 Object detection of apple, banana, and orange

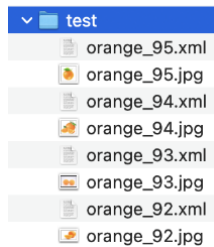
2024 年春季学期

组队说明：2-4 人组队，共享得分，占期末总评 20 分。注：在 5 月 14 日前派组员在企业微信跟助教“吴世伟”或“陈河旭”说明小组成员名单，助教需在对分易上帮忙进行分组。

提交时间：本次作业于 2024.06.17 (周一)晚上 12 点前提交，无合理理由逾期者按逾期天数每天扣 20%（即周二晚 23:59 前算一天，周三晚 23:59 前算两天...）。

提交方式：对分易作业“上机课小组大作业”提交，提交的压缩包命名为“组员 1 名字_组员 2 名字_组员 3 名字_组员 4 名字.zip”（例如：吴世伟_陈河旭_范昊翔_李国文.zip），压缩包应需包含以下文件：

- 任务代码中配来的原始的 test 文件夹



- 一个命名为“大作业报告.pdf”的文件，格式要求参考本次大作业附的“大作业报告.docx”
- 一个命名为“label_data”的文件夹，包含 20 张你们小组标注的图片和对应 .xml 文件
- 一个命名为“frontend”的文件夹，里面包含在浏览器运行你前端交互的代码文件
- 一个 ReadMe.txt 文件，上面说明你们的压缩包里的各个文件(夹)分别是做什么用的，以及怎么样配置一些关键的环境，以运行你们的 predict.py
- 你们小组撰写的 predict.py 文件，该文件在执行 python predict.py 之后，需要能调用你们小组训的模型在 test 文件夹中做测试，并做出以下 metric 的输出（参考如下格式）：

Class	IoU	Precisoin	Recall	Mean Average Precision (mAP50)
All	xxx	xxx	xxx	xxx
Apple	xxx	xxx	xxx	xxx
Banana	xxx	xxx	xxx	xxx
Orange	xxx	xxx	xxx	xxx

- 注：关于上面各 metric 的说明，参考 <https://labelyourdata.com/articles/object-detection-metrics>
- 命名为“运行截图”的文件夹，里面包括：
 - 在你们的电脑上运行 predict.py 的输出截图 predict.png
 - 能反映你们训练模型或验证模型过程的任何一张或多张图片，做好相应命名。
- 其它你们小组作业中产生的代码文件或文件夹，注意在 ReadMe.txt 文件中做好说明
- 注：如果你的一些模型或依赖包很大（比如大于 100M），请不要打包上传到对分易，而是自行上传到你们的网盘（如百度网盘），并把相应链接在 ReadMe.txt 文件中附上。

任务

水果目标检测是一项经典的机器学习任务，针对此领域已有丰富的在线资源和代码示例。在本次团队大作业中，我们的目标是通过给定图片作为输入，实现对图片中的苹果(Apple)、香蕉(Banana)和橙子(Orange)的识别和定位。任务可以分为以下四部分：

1. 数据标注：对图片中的苹果、香蕉和橙子利用框框进行标注，具体见下面图 1。
2. 模型训练：训练模型实现对于图片中的水果识别和定位。
3. 前端开发：实现一个用户友好的前端界面，允许用户上传图片进行实时测试。
4. 课堂演示：准备在课堂上对项目进行详细演示，包括模型训练、结果展示及技术解答。

我们使用这个 Fruit Images for Object Detection 数据集中的 test 文件夹中的 60 张图片作为测试集，请到以下链接查看说明和下载数据集：

<https://www.kaggle.com/datasets/mbkinaci/fruit-images-for-object-detection/data>

具体要求和评分说明如下：

一 . 数据收集与标注（需写入“大作业报告.pdf”，总分 5 分）

请收集 20 张 jpg 格式图片，包括 5 张只包含苹果的，5 张只包含香蕉的，5 张只包含橙子的，5 张包含两种或两种以上(mixed)的水果的，请用英文和数字命名图片文件，采用与原始的 **test 文件夹** 中的样式进行标注，即每张图片配一个同名的.xml 文件，其中 `<name>xxx</name>` 中的 xxx 为 apple, banana, orange 三类，参考原数据集提供的标注工具：<https://github.com/HumanSignal/labellmg>

提示和注意事项

- 这 20 张图片，可以是自己拍摄的，也可以是网上找的。
- 20 张图片和对应的 .xml 文件请保存在命名为“label_data”的文件夹中。
- 你们不一定会用这 20 张图片进行训练，我们一方面是想锻炼你们数据标注的能力，另一方面，有趣的是，我们会从各组提交的图片中抽取一些图片，作为大家在课上现场演示时的输入图片，因此，尽量别搞一些太复杂的图片为难你的同学哟，并且，尽量标注准确哟。
- 如有需要，你可以收集和标注更多图片用于你们的模型训练

评分说明

- 提交“label_data”的文件夹的图片即对应的 .xml 文件完整合理，得 3 分，如有一类水果（苹果，香蕉，橙子，mixed）图片不完整或标注错误，各扣一分，扣完为止
- 提交的“大作业报告.pdf”中有适当的图片来源和标注样例的说明，得 2 分，如无适当说明，酌情扣分

二 . 水果检测模型（需写入“大作业报告.pdf”，总分 5 分）

请提供一个水果检测模型，能在原始 test 文件夹中的 60 张图片上进行测试（我们使用这个 Fruit Images for Object Detection 数据集中的 test 文件夹中的 60 张图片作为测试集，请到以下链接查看说明和下载数据集：

<https://www.kaggle.com/datasets/mbkinaci/fruit-images-for-object-detection/data>）。

你们可以从头开始训练一个模型，可以 fine-tune 一个已有的模型，或者尝试现有（大模型）。你们需要写一个 predict.py 并运行，在 test 文件夹中做测试，并得到上面提交说明

中所指出的 metric 的输出，截图保存为 **predict.png** 提交。

提示和注意事项

- 如有需要，你们可以找更多的现有的水果香蕉橙子数据集做训练，哪怕那些数据集标注时用的不是 .xml 文件，但是，你们得确保写的 predict.py 能在原始的 test 文件夹中做测试你的模型。
- 不限制使用的模型，但建议不要搞太复杂的，因为很抱歉，我们课程不能提供计算资源以供大家训练。
- 如果你们调用大模型，需在实验报告中提供完整的 prompt，并且小组自行承担调用 api 所需的费用。

评分说明

- 在 test 文件夹中做测试，得到的 All, Apple, Banana, Orange 的 $\text{IoU} \geq 0.5$, $\text{precision} \geq 0.7$, $\text{recall} \geq 0.7$, $\text{mAP} \geq 0.7$ ，共 $4 \times 4 = 16$ 个值（需在你提交的 **predict.png** 文件中显示），如果有 12 个或 12 个以上值满足上面的要求(e.g., ≥ 0.7)，则得 3 分，如果有 8-11 个值满足上面的要求，得 2 分，如果有 4-7 个值满足要求，得 1 分，如果小于 4 个值满足要求，得 0 分
- 提交的“大作业报告.pdf”中有适当的模型说明、训练模型或验证模型过程说明，以及训练数据集说明（如有用训练数据集），得 2 分，如无适当说明，酌情扣分

三．水果检测前端 demo（需写入“大作业报告.pdf”，总分 5 分）

请提供一个简单的前端 web 界面，提供以下基本交互： 用户可点击上传一张水果图片，该前端调用你们训练的模型，并在前端显示出框选出的水果图片，效果如下所示（当然，我们只有苹果、香蕉、橙子三类）：

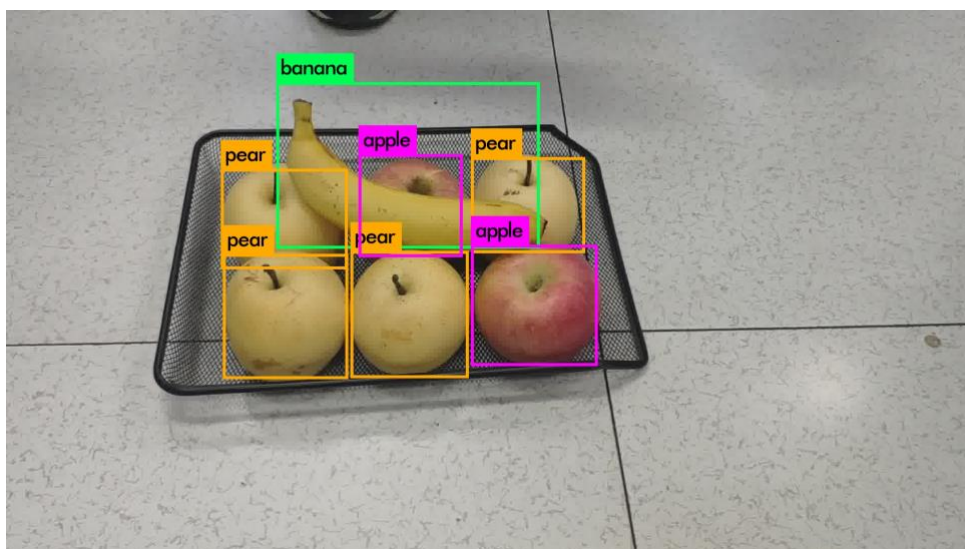


图 1 实验效果图

提示和注意事项

- 完成上面的基本交互和显示功能就行，但是，我们鼓励相对好看一点的前端界面，因为你们将在课上演示。
- 注意，你们得想办法把框框和标签画出来
- 注意，你们的前端应该能应对和现实任意尺寸和大小输入的图片（由各个小组提供）

评分说明

- 提交的命名为“**frontend**”的文件夹，有可执行的前端代码文件，得 3 分，注意：助教会抽查一些小组，请不要冒险欺骗助教。
- 提交的“**大作业报告.pdf**”中有适当的前端实现说明和实现后的前端截图，得 2 分，如无适当说明，酌情扣分

四．课堂演示（无需写入“大作业报告.pdf”，总分 5 分）

现在是 show time，我们计划在某一节课上（理论课或实验课），请各个小组用你们小组成员的电脑，演示你们的前端 demo，不同的是，我们会随机抽取各个小组提交的标注图片，每组测试 8 张。

评分说明

- 测试的 8 张图片中，能正确框出 6 张图片中的所有正确水果，即得 5 分，如果只能正确识别 5 张，得 4 分，只正确识别 4 张，得 3 分，以此类推。