看整张图，还是找图里的东西?

图像分类与目标检测学生阅读手册

适合小学高年级和初中学生。阅读约 10 分钟，也可以跟着老师在飞令 FeeLink 2.24 里动手试一试。

1. 同一张照片，两种问题

想象你给教室拍了一张照片。照片里有两把椅子、一个书包，还有一张桌子。下面两个问题都能交给图像识别模型：

你想问什么?	选哪种任务?	可能得到的答案
这张照片拍的是教室、走廊还是操场?	**图像分类**	整张照片最像「教室」
照片中哪些地方有椅子和书包?	**目标检测**	给两把椅子和一个书包分别画框，并写出它们的类别

图像分类给整张图一个答案。**目标检测**给画面中的每个已找到的目标一个位置和类别。同一类可以出现好几次，每个找到的目标各有一个框。只要知道整张图是什么，可以用分类。想找位置，或想数出画面里找到了几个目标，可以用检测。计数时也要留意漏框和多框。 Ultralytics 图像分类 · Ultralytics 目标检测

同一张照片在飞令里的两种结果

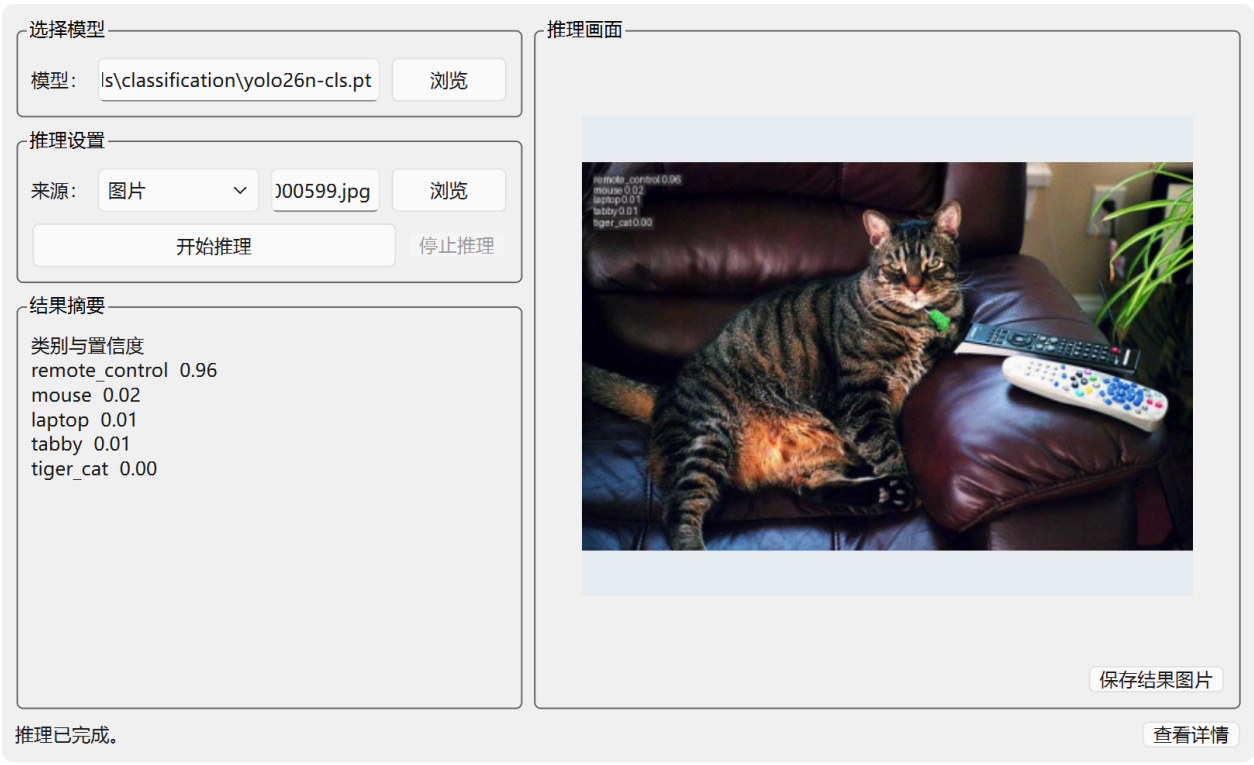


图 1　分类模型给整张照片打分。这次排名第一的是 `remote_control`（遥控器），界面也列出了其他候选类别。

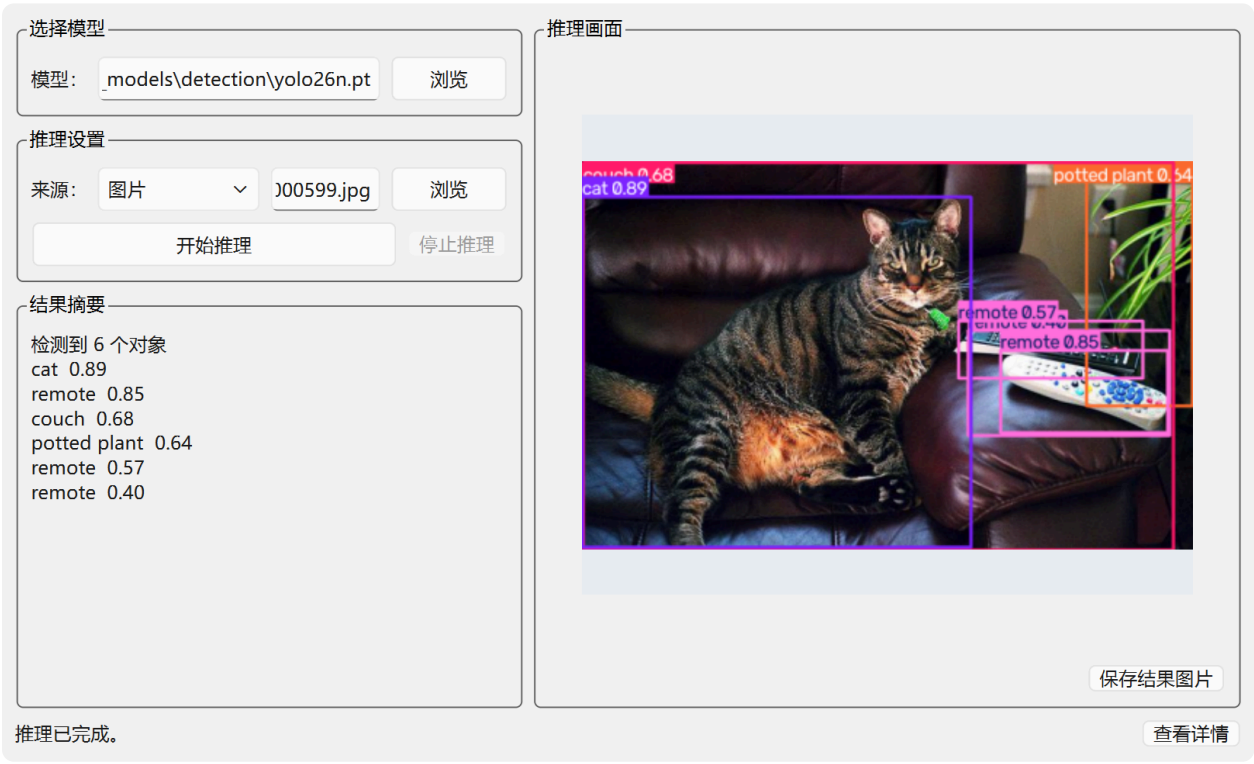


图 2　换用检测模型后，同一张照片里出现了多个框。标着 `cat` 的框圈出了猫。看看遥控器附近有几个框，再和原图里的物体数量比一比。

> 先想一想：要认出纸上单独写的数字是几，该选哪种任务？如果一张纸上写了好几个数字，还想知道每个数字在哪里呢？

2. 飞令、Ultralytics 和 YOLO 分别是什么?

飞令（FeeLink）是你在电脑上操作的软件。你可以用它整理图片、给物体画框、训练模型，再把新图交给模型试一试。

Ultralytics 是飞令用到的 Python 工具库。你可以把工具库理解成程序员借用的工具箱。在飞令里点击「开始训练」或「开始推理」时，软件会调用这个工具箱里的功能。跟着飞令练习时，直接操作界面即可。

YOLO 是这里用到的一系列图像识别模型。把模型想象成一个做过练习的答题器：训练时，它看图片和已知答案，调整自己判断的方法。推理时，它看一张新图，给出自己的预测。飞令随包提供已经练过的模型文件，可以直接拿合适的模型试用，也可以用它作为起点来训练自己的模型。 Ultralytics 图像分类 · Ultralytics 目标检测

3. 图像分类：认出一个手写数字

假设我们想让模型认出单独写在一张纸上的 `3` 或 `7`。

1. **收集图片。** 请几位同学写数字，每张图只保留一个数字。用不同的笔或写法，让模型见过不同样子。
2. **按答案分文件夹。** 写着 `3` 的图片放进 `3/`，写着 `7` 的图片放进 `7/`。文件夹名就是整张图的答案。飞令选择「图像分类」时会关闭画框入口。
3. **划分数据、训练模型。** 在「制作数据集」里检查格式并划分。训练集和验证集都按类别分文件夹，例如 `train/3/`、`train/7/`、`val/3/`、`val/7/`。在「训练模型」里检查数据后开始训练。
4. **换张新图试一试。** 选一张训练时没有用过的手写数字图，用训练好的模型推理。先记下它猜的是几，再和你看到的数字比较。

飞令自带已经划分好的手写数字参考数据，包含 `0` 到 `9`。如果老师让你用随包数据，选「图像分类」，直接在「训练模型」里选 `reference_datasets/classification/`，然后用其 `test/` 目录中的图片推理。 Ultralytics 分类数据目录

4. 目标检测：找到桌上的杯子和瓶子

假设我们想让模型找到桌面照片里的杯子和瓶子。

1. **收集图片。** 拍几种不同摆法。有时一张图有两只杯子，有时杯子和瓶子一起出现，也可以有一张桌面上两种目标都没有的照片。
2. **给每个目标画框。** 在飞令选「目标检测」，用「标注图像」给每个杯子或瓶子画一个贴住物体的矩形框。提前约定 `0` 是杯子、`1` 是瓶子。如果一张图有两只杯子和一个瓶子，就画三个框，其中两个框的类别编号都是 `0`。
3. **保存标注、制作数据集。** 图片旁会有同名的标签文件，例如 `desk01.jpg` 对应 `desk01.txt`。标签记录每个框的位置、大小和类别。划分数据集后，图片与它的标签要一起进入同一组。生成检测数据配置文件（YAML），它告诉模型去哪里找图片，以及编号对应什么类别。

4. **训练后用新图推理。** 在训练页检查数据并训练。拿一张新桌面照来试，看看杯子和瓶子有没有被找到，框的位置对不对。结果可能出现漏框、类别认错或多画框，逐个对照原图检查。

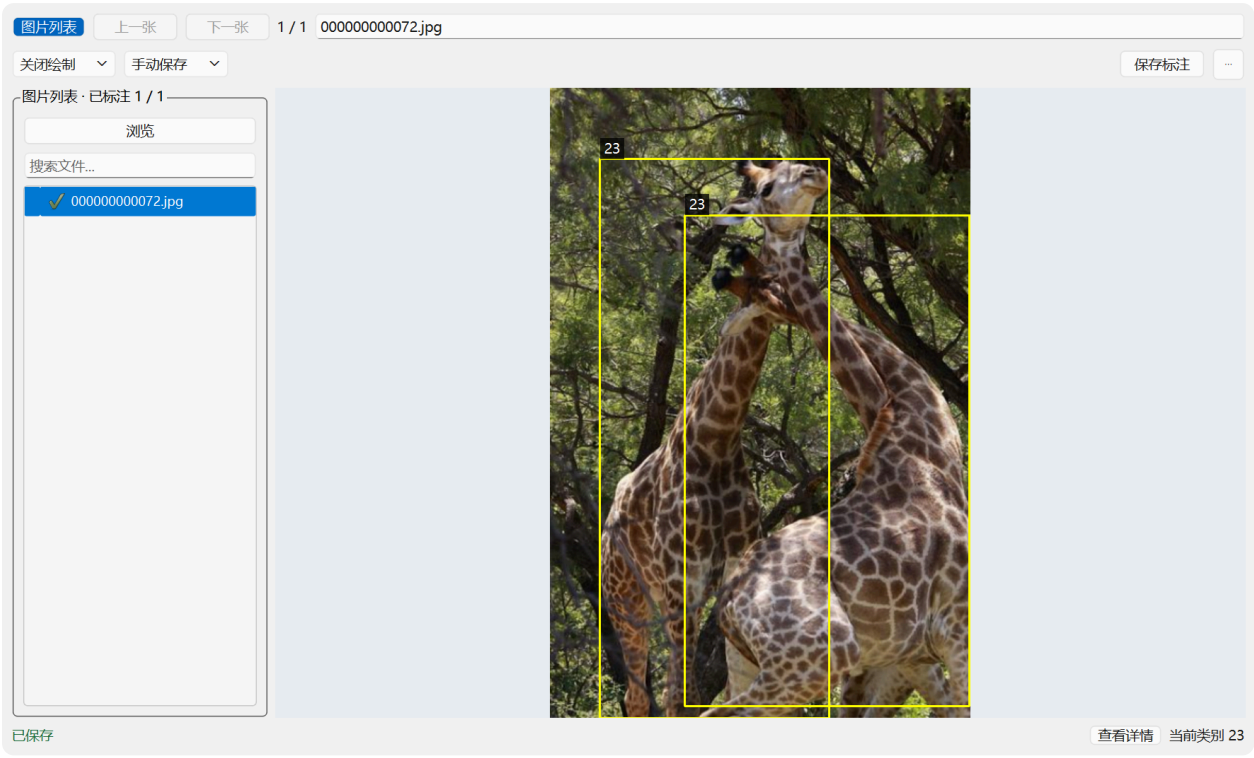


图 3　训练前由人画的框叫「标注」。这张图给两只长颈鹿各画了一个框。两个框都标着 23，这是参考数据给长颈鹿用的类别编号。图 2 的框则是模型自己预测出来的。

飞令随包的 `reference_datasets/detection/` 收录了筛选过的动物、交通工具和生活用品照片，带有标注和配置文件。它适合练习读框和看结果。训练自己的杯子和瓶子模型时，使用你们为这个任务收集并标注的图片。 Ultralytics 检测数据格式

5. 模型怎样练习？我们怎样检查？

数据分组	可以把它想成	用来做什么
训练集（`train`）	带答案的练习题	模型根据图片和答案调整自己的判断
验证集（`val`）	学习过程中的小测	观察训练过程中模型在另一批图片上的表现
测试集（`test`）	最后留出的测验	训练结束后，用独立图片检查模型能否处理新例子

把同一张照片的原图、裁剪版和翻转版放在同一组。如果模型练习时见过一张照片，后来又用它的翻转版来考试，测出来的成绩可能偏高。 Ultralytics 分类数据目录 · Ultralytics 检测数据格式

模型给一次预测打的分叫**置信度**。例如「瓶子 0.90」是模型对这一次预测给出的分数，它仍可能认错。**准确率**要把一批分类图片的预测与答案比较。例如 10 张数字图里猜对 8 张，这 10 张图上的准确率就是 80%。检测时还要检查框的位置，看看框住的是不是目标本身。 Ultralytics 图像分类 · Ultralytics 目标检测

动脑题：模型可能学错了什么?

假如你拍的所有杯子都放在白桌上，所有瓶子都放在黑桌上。现在把杯子换到黑桌上，模型会怎样猜?

先试一次，把结果记下来。模型可能把桌面颜色当成线索。你可以让杯子和瓶子都在不同颜色的桌面上出现，重新收集数据，再用另外拍的照片检查结果。

6. 在飞令里走一遍

你手上有……	从这里开始
自己拍的图片	选任务，按类别分文件夹或画框，再制作数据集
随包已划分的参考数据	选对应任务，从「训练模型」开始
老师给的模型文件	从「部署模型」开始，选图片并点击「开始推理」

训练成功后可以点「用此模型推理」。也可以到「部署模型」里打开训练得到的 `best.pt`。课堂练习结束时，保存一张结果图，并写下：我用的是哪种任务？模型答对了什么？它有没有猜错或漏掉什么?

7. 三道小测

1. 把每张照片分到「猫」或「狗」文件夹，应该用什么任务?
2. 一张图里有两只猫，你想知道两只猫各在哪里。标注这张图时应该画几个框?
3. 同一张图的原版在训练集，翻转版在测试集。怎样让这个测试更公平?

参考答案

1. 图像分类，因为你要给整张图一个类别。
2. 目标检测，画两个框，都标为「猫」。
3. 把原版和翻转版放在同一组，再用另外拍的照片做测试。

使用反馈

遇到软件操作问题，可在「帮助 → 使用反馈」复制邮箱或网站网址。邮箱为 yujunzhan@163.com。
飞令网站反馈