

(1) 赛题任务

本次比赛要求参赛选手构建一个高精度的花卉识别模型，能够准确识别和分类不同种类的花卉。

- 识别100种花卉，包含稀有品种和形态相似的花卉。

(2) 数据使用规则

- 本次比赛允许参赛选手使用主办方提供的训练数据以及指定的公开数据集。
- 允许使用数据增强技术，但不允许使用外部标注数据。
- 可以使用预训练模型，但需在技术报告中详细说明。

2. 数据集描述

(1) 主办方提供数据集：

- 训练集：包含100个花卉类别，每个类别 100-150张高质量图片，用于模型训练或者调优。
- 测试集：包含100个花卉类别，每个类别50张图片，不提供下载，保存在服务器后台用于最终评测

(2) 数据格式：

- 图片格式： JPG/PNG
- 图片尺寸：统一调整为 600×600
- 标注格式： CSV文件，包含图片路径和类别标签

(3) 测试集与项目结构：

训练集结构：

```
filename,category_id,chinese_name,english_name
img_000051.jpg,164,紫叶竹节秋海棠（紫竹梅）,Tradescantia pallida
img_000052.jpg,164,紫叶竹节秋海棠（紫竹梅）,Tradescantia pallida
...
```

项目结构：

```
Flower_Classification:
```

```
  |- test_dataset
  |  |- img_001.jpg
  |- results
  |  |- model.py
```

```
|- requirements.txt  
|- utils.py  
+- data  
|   |- train_labels.csv  
|   |- train  
|       |- img_000051.jpg  
+- code  
|   |- predict.py  
|   |- train.py
```

3. 模型输出文件格式

选手模型需要生成一个.csv格式的结果文件，编码为UTF-8，格式如下：

Img_name	predicted_class	confidence
Img_001.jpg	5	0.95
...