

实验题目: 图像分类标注与模型验证

实验目标

1. 理解文本标注的基本概念和流程。
2. 理解图像分类标注的基本概念和流程。
3. 理解标注数据对模型训练的重要性
4. 掌握图像标注工具的使用方法。
5. 使用标注工具对图像数据进行标注并分析标注质量。

实验内容

1. 数据准备
 - 将本次实验的数据集文件中的60张图片进行标注。
 - 图像标注应包含三个类别(猫、狗、鸟)。
2. 图像标注
 - 标注工具: 使用标注工具(如Label Studio、[Labellmg](#))为每张图片添加分类标签。

Label Studio:

安装: `pip install label-studio`

启动: `label-studio start`

访问: 打开浏览器, 输入 <http://localhost:8080>。

- 标注格式: 导出标注结果为CSV文件。

示例:

```
image_path, label
cat_001.jpg, cat
dog_005.jpg, dog
...
```

3. 模型训练

- 数据预处理:
 - i. 将图片统一调整为 224*224 像素, 并进行归一化(像素值缩放到 $[0, 1]$)。
 - ii. 划分数据集: 训练集(80%)与测试集(20%)。
- 模型选择(任选一个模型, 可自定):
 - i. 使用PyTorch或TensorFlow搭建一个简单CNN模型(参考结构: 2层卷积 + 2层全连接)。
 - ii. 或直接使用预训练模型(如ResNet-18)进行迁移学习。
- 模型训练:
 - i. 设置超参数: 学习率 0.001, 优化器 Adam, 训练轮数 10。
 - ii. 保存训练过程中的损失和准确率曲线。

4. 模型评估

- 评估指标: 准确率(Precision)、召回率(Recall)、F1分数(F1-Score)。

- 结果分析:检查标注错误的样本(如标签混淆、图片模糊),讨论其对模型的影响并提出改进标注质量的方案(如多人标注交叉验证)。

实验要求

1. 提交材料:

- 提交标注文件(CSV)、训练代码、模型评估结果。
- 实验报告,重命名为“**学号姓名.docx**”。
- 截止日期:2025年3月30日23:55。

2. 实验报告内容:

- 数据集描述:数据来源、数据量、图片类型。
- 标注过程:标注工具、标注标准、标注一致性检查方法。
- 模型训练:模型选择、超参数设置、训练过程。
- 结果分析:评估指标、模型表现、错误案例分析。