

**课程报告**

**报告名称： 专业综合实践**

**实验项目： 猫狗蝴蝶分类项目**

**班 级：21级数据科学与大数据技术1班**

**姓 名： 杨嘉杭**

**学 号： 419470146**

提交日期：2024年11月22日

**目录**

[一、项目背景与目的 1](#_Toc17118)

[背景： 1](#_Toc26341)

[目的： 1](#_Toc18700)

[二、数据预处理 2](#_Toc11596)

[数据集描述： 2](#_Toc16417)

[预处理步骤： 2](#_Toc27095)

[归一化处理： 3](#_Toc9028)

[划分数据集： 3](#_Toc31680)

[三、模型构建 3](#_Toc10934)

[模型选择： 3](#_Toc24710)

[模型架构： 4](#_Toc29488)

[描述激活函数、损失函数和优化器的选择 5](#_Toc27214)

[四、模型评估 5](#_Toc21841)

[评估指标： 5](#_Toc18978)

[1. 准确率（Accuracy） 5](#_Toc5460)

[2. 召回率（Recall） 5](#_Toc5440)

[3. F1 分数 6](#_Toc5543)

[评估方法： 6](#_Toc26079)

[五、结果分析与优化 8](#_Toc19233)

[结果分析： 8](#_Toc31502)

[模型优化： 9](#_Toc26590)

[附录 9](#_Toc6298)

# 基于深度学习的图像识别系统：猫狗蝴蝶分类项目

## 一、项目背景与目的

### 背景：

在人工智能和深度学习快速发展的今天，图像分类技术已被广泛应用于各行各业，例如医疗诊断、智能监控、自动驾驶等场景。图像分类不仅可以高效地从海量图像中提取有价值的信息，还为图像识别和理解奠定了基础。

猫、狗、蝴蝶是日常生活中常见的动物种类，其外观差异明显，但也存在一些复杂场景（如光线变化、角度不同、遮挡）会对分类算法提出挑战。通过对猫、狗、蝴蝶进行图像分类的研究，不仅可以验证深度学习模型的效果，还可以为类似的多分类任务提供技术支持。

### 目的：

**构建一个高效的图像分类系统**：

基于深度学习技术，设计并训练一个分类模型，能够自动识别猫、狗和蝴蝶三种类别。

**验证深度学习模型在小型数据集上的性能**：

通过卷积神经网络（CNN）模型，探索图像分类的效果，评估模型在训练集和验证集上的表现。

**提升模型泛化能力**：

应用数据增强和正则化等技术，降低过拟合风险，提升模型在未见数据上的表现。

**探索技术应用场景**：

为相关领域（如动物保护、宠物分类管理和自然环境监测）提供潜在的技术支持与解决方案。

**学习和优化深度学习技术**：

通过模型搭建和训练，熟悉深度学习模型的构建、优化和评估流程，为未来的图像处理任务打下基础。

## 二、数据预处理

### 数据集描述：

数据集包含三类图片：猫、狗 和 蝴蝶。

数据存储结构：

每个类别对应一个文件夹，文件夹名称为类别名，例如 cat/, dog/, butterfly/。

数据来源是本地目录或公开数据集。

每类图片的分辨率可能不一致，需在预处理时统一大小。

### 预处理步骤：

使用 **Torchvision.transforms** 模块进行数据预处理，主要步骤：

**调整大小**：将所有图像统一调整为指定大小（如 128x128）。

**数据增强**：

随机水平翻转（RandomHorizontalFlip）。

随机裁剪（RandomCrop）。

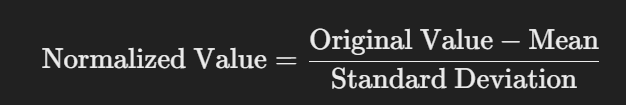
**转换为张量**：将图像从 PIL 格式转换为张量格式。

**归一化**：对像素值进行归一化处理。

### 归一化处理：

**方法**：对像素值进行归一化，使其范围从 [0, 255] 转换为 [-1, 1]，提高训练稳定性。

**计算公式**：



**均值和标准差**：

mean = [0.5, 0.5, 0.5]：针对 RGB 三通道的均值。

std = [0.5, 0.5, 0.5]：针对 RGB 三通道的标准差。

### 划分数据集：

**数据集划分比例**：

**训练集**：70% 用于模型训练。

**验证集**：20% 用于调参和模型验证。

**测试集**：10% 用于模型最终评估。

**方法：**

使用 torchvision.datasets.ImageFolder 加载数据集。

利用 torch.utils.data.random\_split 将数据集划分为训练集、验证集和测试集。

## 三、模型构建

### 模型选择：

CNN 的自定义卷积神经网络，用于猫、狗和蝴蝶分类任务。该模型结构较为简单，适合中小型数据集的分类任务。

### 模型架构：

**卷积层部分**

**输入通道数**：3（RGB 图像）。

包含三个卷积层，逐步增加通道数：

第1层：输入通道为 3，输出通道为 32，卷积核大小为 3×33 \times 33×3，步长为 1，填充为 1。

第2层：输入通道为 32，输出通道为 64，卷积核大小为 3×33 \times 33×3，步长为 1，填充为 1。

第3层：输入通道为 64，输出通道为 128，卷积核大小为 3×33 \times 33×3，步长为 1，填充为 1。

每个卷积层后接：

**ReLU 激活函数**：增加非线性特性。

**最大池化层**：池化核大小为 2×22 \times 22×2，步长为 2，用于降维。

**全连接层部分**

**Flatten 层**：将卷积层输出展平。

两个全连接层：

第1层：输入维度 128×16×16=32,768128 \times 16 \times 16 = 32,768128×16×16=32,768（假设输入图像大小为 128×128128 \times 128128×128），输出维度为 256。

第2层：输入维度为 256，输出维度为类别数（3）。

激活函数：

使用 ReLU 激活函数。

在第1个全连接层后添加 **Dropout 层**（比例为 0.5），防止过拟合。

### 描述激活函数、损失函数和优化器的选择

**激活函数**：

卷积层和全连接层使用 ReLU 激活函数。

**损失函数**：

使用 **交叉熵损失函数**（nn.CrossEntropyLoss），适合分类任务。

**优化器**：

使用 Adam 优化器（torch.optim.Adam），学习率为 0.001。

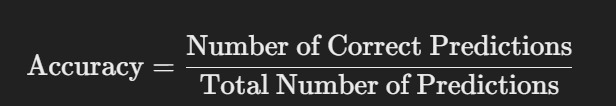
## 四、模型评估

## **评估指标：**

### 1. 准确率（Accuracy）

描述：模型预测正确的样本比例。

计算公式：

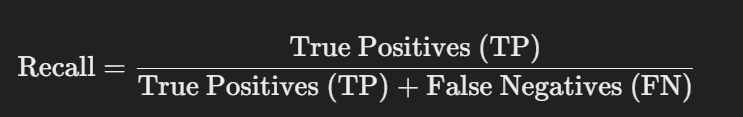


适用场景：当各类别的样本数大致相同时，准确率是合适的指标。

### 2. 召回率（Recall）

描述：模型正确预测为某类别的样本比例。

计算公式：

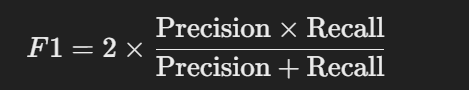


适用场景：当漏报（FN）代价较高时，例如灾害预测。

### 3. F1 分数

描述：精确率和召回率的调和平均值，用于综合评估模型性能。

计算公式：



适用场景：当需要平衡误报和漏报的影响时。

## **评估方法：**

**交叉验证**：

描述：将数据集分为若干折（folds），交替使用其中一折作为验证集，其余折作为训练集。

作用：减少模型对特定数据分割的依赖，提高模型泛化能力。

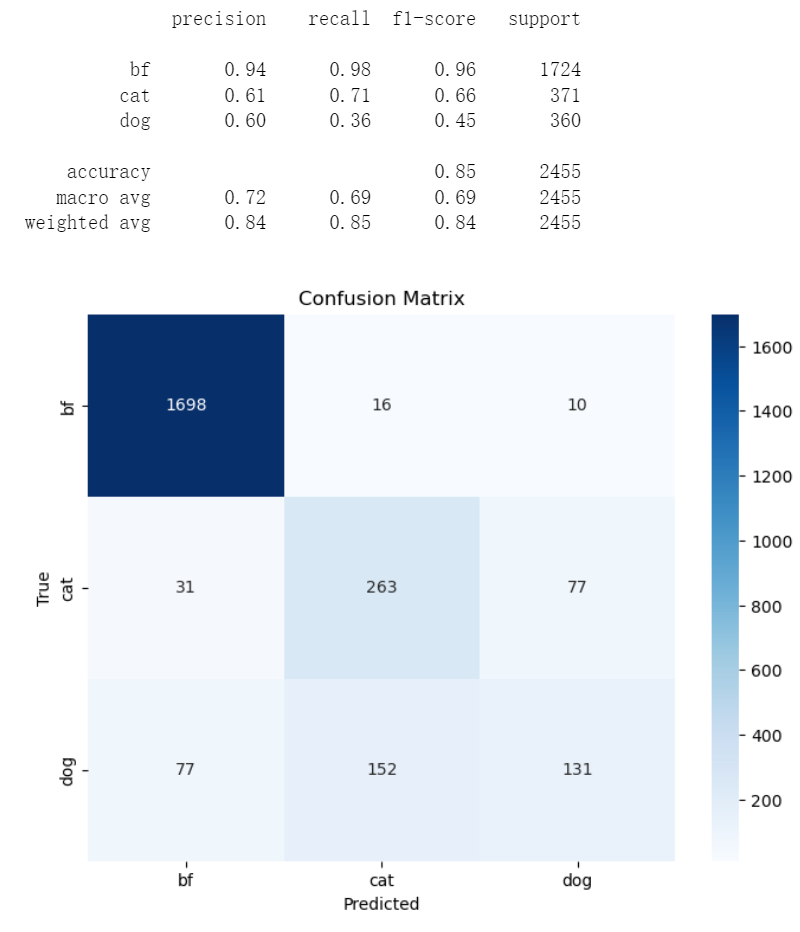
**混淆矩阵**：

描述：显示模型预测与实际值的对比情况。

矩阵结构：

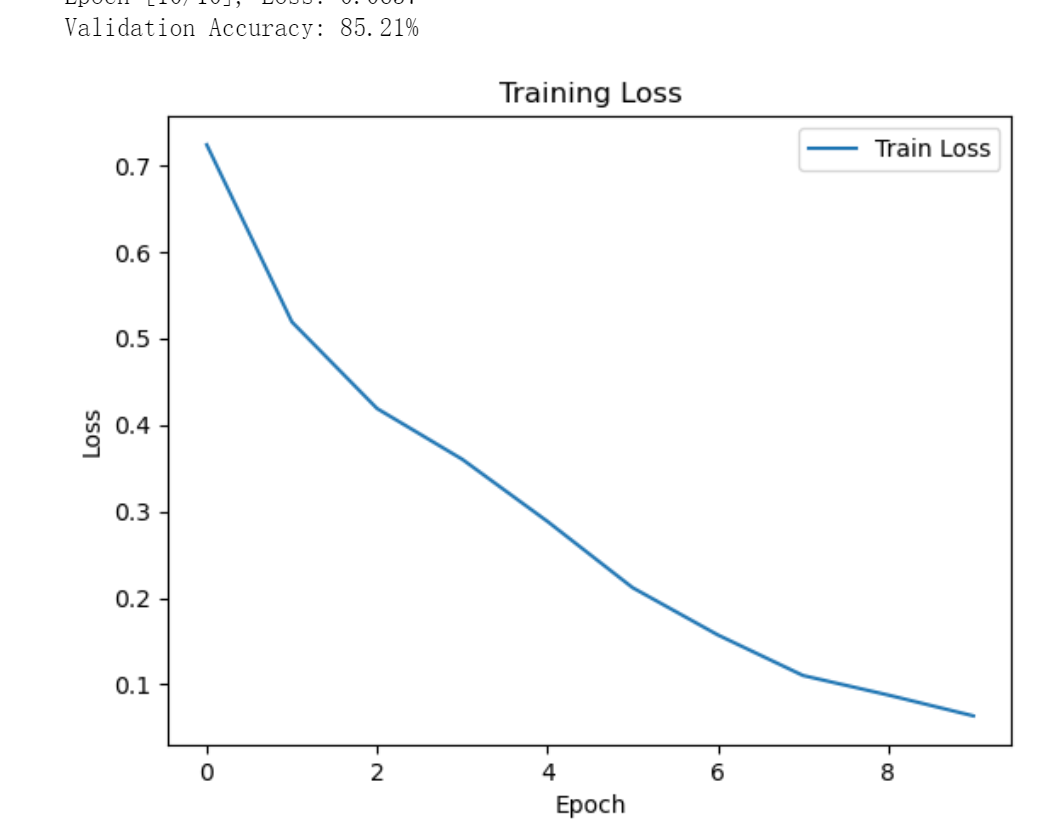
行表示实际类别。

列表示预测类别。

主对角线上的值为预测正确的样本数。

## 五、结果分析与优化

### 结果分析：



**收敛性良好**：

损失曲线平稳下降，未出现波动或上升趋势，说明训练过程较为平稳，模型逐步学到了训练数据的特征。

**训练充分**：

损失值在 10 个 Epoch 内显著下降，表明训练过程有效且模型对训练数据有较好的拟合能力。

当前损失值接近 0.1，表明模型对训练数据的预测误差较小。

**验证准确率较高**：

验证集准确率为 **85.21%**，表明模型对未见数据的预测效果较好，泛化能力较强。

### 模型优化：

**数据增强**：

增加数据增强策略（如随机旋转、裁剪、翻转等），提升模型的泛化能力。

**模型结构优化**：

如果准确率未达预期，可尝试引入预训练模型（如 ResNet 或 EfficientNet）以提升性能。

**超参数调整**：

适当调整学习率、批量大小等超参数，观察对训练曲线的影响。

**增加训练数据**：

如果可能，增加更多的训练数据，特别是对于表现较差的类别，帮助模型学习更全面的特征。

## 附录

github项目链接：https://github.com/Yannel-Y/-.git

附带仓库目录结构截图

