

**课程报告**

**报告名称： 专业综合实践**

**实验项目： 狗鸭鹿分类项目**

**班 级：21级数据科学与大数据技术1班**

**姓 名： 庄严**

**学 号： 421470166**

提交日期：2024年11月22日

目录

**[一、项目背景与目的 1](#_Toc20227)**

[背景： 1](#_Toc18689)

**[二、数据预处理 3](#_Toc8817)**

[数据集描述： 3](#_Toc26583)

**[三、模型构建 4](#_Toc27320)**

[模型选择： 4](#_Toc20556)

**[四、模型评估 5](#_Toc24220)**

[1. 准确率（Accuracy） 5](#_Toc6520)

[2. 召回率（Recall） 5](#_Toc3760)

[3. F1 分数 5](#_Toc30740)

**[评估方法： 6](#_Toc29042)**

**[五、结果分析与优化 7](#_Toc29350)**

**[结果分析： 7](#_Toc15691)**

**[附录 8](#_Toc28683)**

# 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：

图像分类是计算机视觉领域中的一项基础且关键任务，它旨在将图像自动归类到预定义的类别中。随着深度学习技术的快速发展，图像分类技术取得了显著进步，尤其是在卷积神经网络（CNN）被广泛应用于该领域之后。图像分类技术的应用范围广泛，包括但不限于安防监控、医疗诊断、自动驾驶以及内容过滤等。近年来，随着大数据和高性能计算资源的可用性增加，以及算法的不断优化，图像分类的准确性和效率都有了大幅提升。此外，随着迁移学习、生成对抗网络（GANs）等新技术的涌现，图像分类技术正朝着更加智能化和精细化的方向发展。这些技术的发展不仅推动了学术界对图像分类问题的研究，也为工业界提供了强大的技术支持。

### 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。

CNN算法：

卷积神经网络（Convolutional Neural Network，CNN）是一种深度学习模型，特别适用于处理图像数据。其主要特点是通过卷积操作捕捉图像中的空间信息，并通过池化操作降低数据维度，从而有效地学习图像的特征和模式。CNN模型的结构包括卷积层、池化层和全连接层。

1. 卷积层：卷积操作利用一组卷积核（也称为滤波器）对输入图像进行特征提取。卷积核在输入图像上滑动并计算局部区域的加权和，从而生成特征图。通过多个卷积核，CNN可以学习不同抽象级别的特征，如边缘、纹理和形状。

2. 池化层：池化操作用于降低特征图的空间维度，减少计算量并增强模型的鲁棒性。常见的池化方式包括最大池化（Max Pooling）和平均池化（Average Pooling），它们分别取局部区域的最大值或平均值作为池化后的值。

3. 全连接层：全连接层接受池化层输出的特征向量，并通过权重矩阵进行线性变换和非线性激活，从而实现对特征的进一步提取和分类。

CNN模型的训练通常包括前向传播和反向传播两个阶段。在前向传播过程中，输入数据经过卷积、池化和全连接层，得到模型的预测输出。通过与真实标签计算损失函数，并利用反向传播算法更新模型参数，使损失函数最小化，从而完成模型训练。

CNN模型在图像处理领域取得了巨大成功，被广泛应用于图像分类、目标检测、语义分割等任务中，成为深度学习领域的重要技术之一。其优势在于能够有效地提取图像中的层次化特征，并具有一定的平移和尺度不变性，适用于处理各种大小和形状的图像数据。

## 二、数据预处理

### 数据集描述：

数据集包含三类图片：狗、鹿 和 鸭子。

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 duck/, dog/, deer/。

### 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：

### 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。

## 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。

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。

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。

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**：

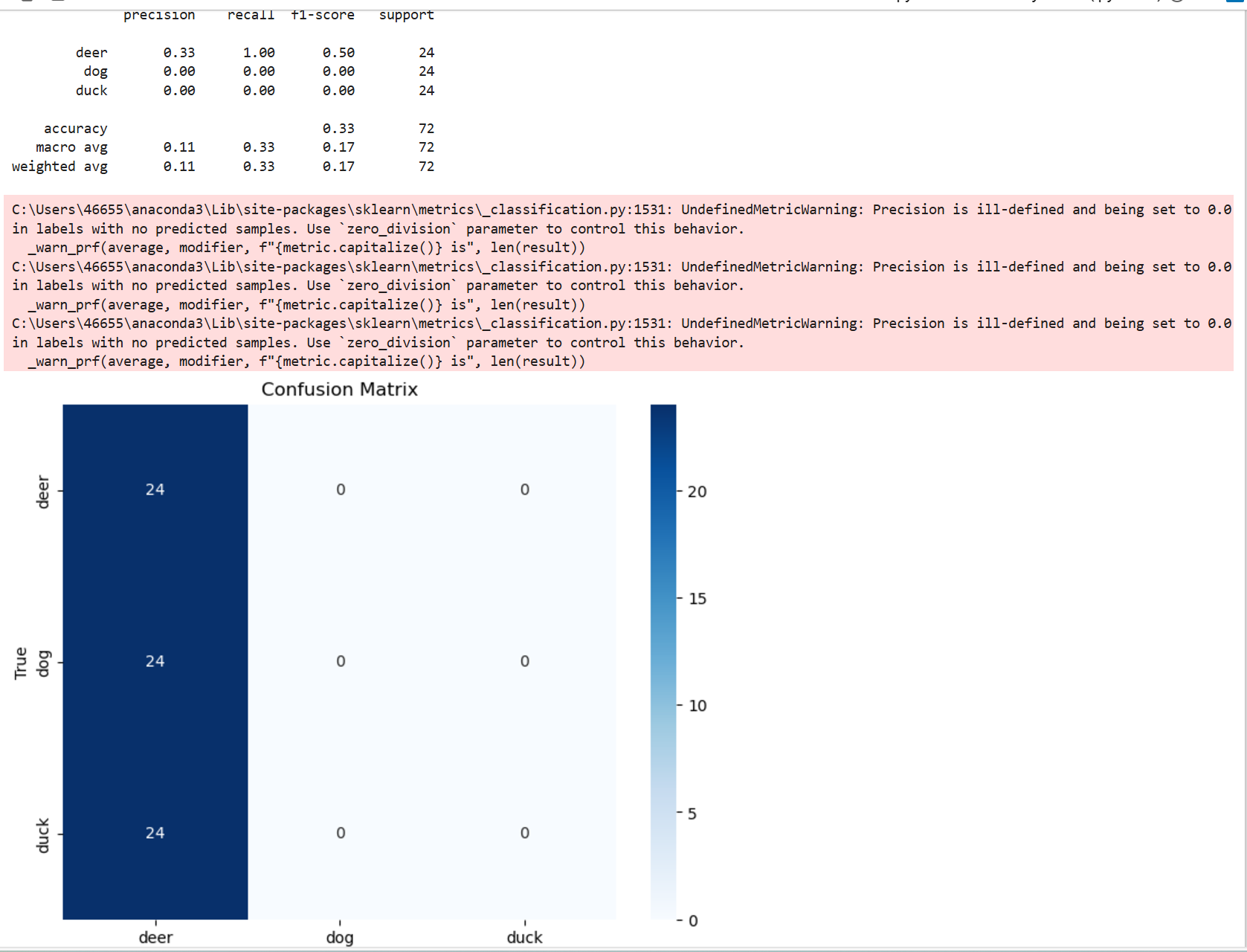
描述：显示模型预测与实际值的对比情况。

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：

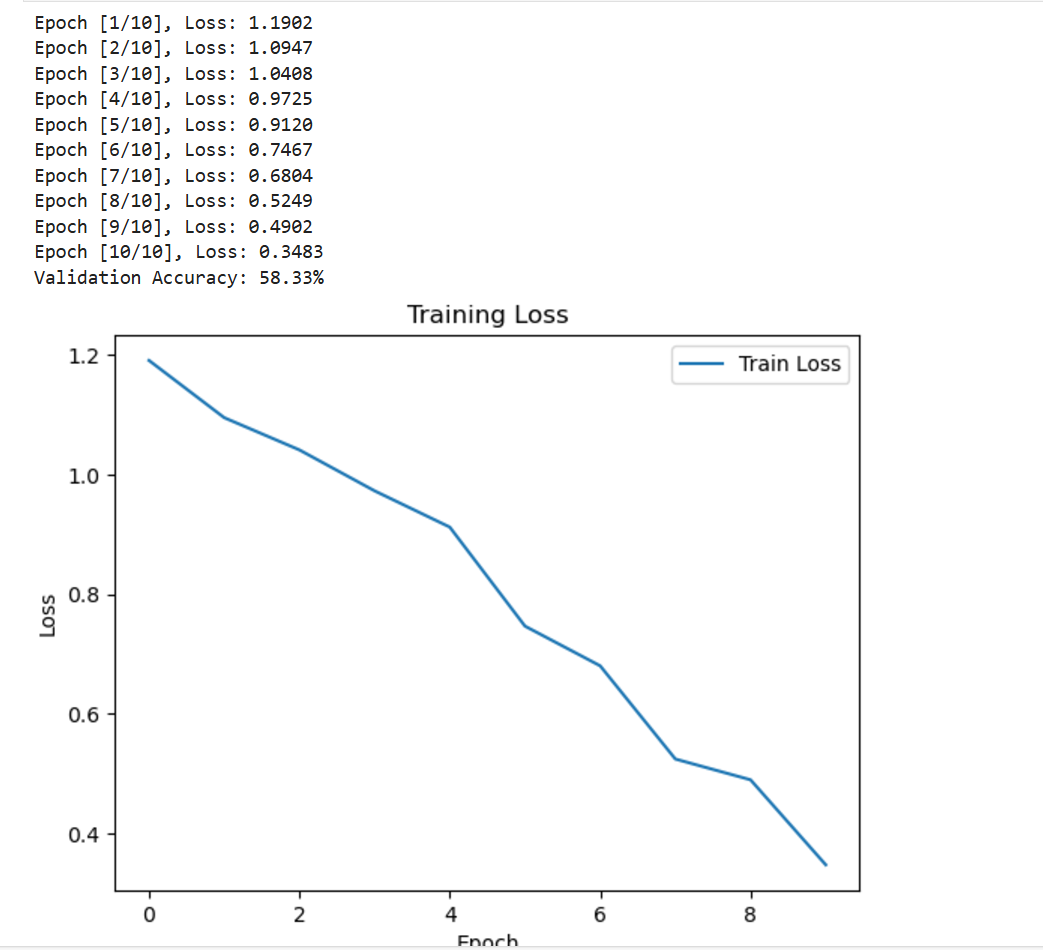
**收敛性良好**：

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。

。

**验证准确率良好**：

验证集准确率为 58.33**%**，表明模型对未见数据的预测效果良好，泛化能力较强。



## 附录

github项目链接：https://github.com/POIZYS/zhuangyan.git

附带仓库目录结构截图

