**图像识别项目报告：猫狗分类问题**

**一、项目背景与目的**

背景

图像识别技术在人工智能领域扮演着至关重要的角色。它不仅能够识别和分类图像，还能够在多种应用场景中提供智能决策支持。猫狗分类问题作为图像识别中的一个经典案例，不仅因其广泛的社会关注度而具有实际应用价值，而且也是研究图像识别技术的一个良好起点。

目的

本项目旨在通过猫狗分类问题，培养学生对图像处理和机器学习的基本理解。通过实践，学生将学习如何使用深度学习模型解决实际问题，并在此过程中提高数据预处理、模型构建、评估和优化的技能。

**二、数据预处理**

数据集描述

本项目所使用的猫狗图像数据集来源于本地存储路径 "data/training\_data/"，包含了大量猫和狗的图像，具有多样性和代表性。这些图像将作为训练和测试深度学习模型的基础。

预处理步骤

为了使模型能够有效地学习和预测，我们对数据集进行了以下预处理步骤：

图像尺寸调整：使用 transforms.Resize((256, 256)) 将所有图像调整至统一的尺寸，以适应模型的输入要求。

归一化处理：使用 transforms.Normalize(mean=(0.5, 0.5, 0.5), std=(0.5, 0.5, 0.5)) 将图像的像素值缩放到0到1的范围内，以提高模型训练的稳定性和效率。

数据增强：通transforms.RandomHorizontalFlip(p=0.3) 和 transforms.RandomVerticalFlip(p=0.3) 方法增加数据的多样性，以增强模型的泛化能力。

划分数据集：将数据集划分为训练集、验证集和测试集，以便于模型的训练和评估。

**三、模型构建**

模型选择

本项目选择了卷积神经网络（CNN）作为深度学习模型，因其在图像识别任务中表现出色。CNN能够自动学习图像的特征表示，无需手动提取特征。

模型架构

模型的架构包括多个卷积层、池化层和全连接层。具体如下：

卷积层：模型包含三个卷积层，每个卷积层后接批量归一化层（nn.BatchNorm2d）和ReLU激活函数，以及最大池化层（nn.MaxPool2d）。

全连接层：在卷积层之后，模型包含两个全连接层（nn.Linear），第一个全连接层将特征图展平并映射到64维空间，第二个全连接层将64维特征映射到10维空间。

输出层：最终通过一个线性层（nn.Linear）将10维特征映射到2维（对应猫狗分类），并使用softmax函数进行分类概率的输出。

**四、模型评估**

评估指标

本项目采用了准确率作为主要的性能指标来评估模型的性能。

评估方法

交叉验证：通过在训练过程中的验证集上评估模型性能，确保模型的泛化能力。

测试集评估：在测试集上评估模型的最终性能，以确定模型在未见数据上的表现。

**五、结果分析与优化**

结果分析

通过对比不同模型的性能，分析其优缺点，并讨论模型在特定类别上的表现差异，以识别模型的潜在问题。

模型优化

根据结果分析，调整模型参数，并尝试不同的网络结构或正则化技术以提高模型的性能。

附录

github项目链接（含代码、数据） 附带仓库目录结构截图

github项目链接：<https://github.com/wighdes/-.git>

仓库目录结构截图：  
