



廣州華商學院
GUANGZHOU HUASHANG COLLEGE

《专业综合实践 II》课程考核

题 目： 基于深度学习的花卉图像识别系统

学 院： 人工智能学院

专 业： 数据科学与大数据技术

年级班别： 21 大数据 1 班

学 号： 421470119

姓 名： 柯敬耀

指导教师： 刘胜

职 称： 无

提交日期： 2024 年 11 月

目录

1. 绪论	1
1.1 研究背景	1
2. 数据描述与预处理	1
2.1 数据描述	1
2.2 数据集的来源和组成	1
2.3 数据集的特点	1
2.4 数据预处理	2
2.4.1 去重处理	2
2.4.2 模糊图像处理	2
2.4.3 图像尺寸调整	3
3. 模型建立与求解	3
3.1 数据准备	3
3.2 支持向量机（SVM）	3
3.2.1 模型假设	3
3.2.2 损失函数和优化器	3
3.2.3 模型训练	4
3.3 多层感知器（MLP）	4
3.3.1 模型假设	4
3.3.2 模型定义	4
3.3.3 损失函数和优化器	4
3.3.4 模型训练	4
3.4 残差网络（ResNet）	4
3.4.1 模型假设	5
3.4.2 模型定义	5
3.4.3 模型微调与训练	5
3.5 模型评估	5
4. 结果展示	5
4.1 SVM 模型评价	5
4.2 MLP 模型评价	6
4.3 ResNet 模型评价	7
5. 结论	8

基于深度学习的图像识别系统：花卉分类项目

1. 绪论

在当今数字化和信息化的时代，图像分类作为计算机视觉领域的一个关键问题，引起了广泛的研究兴趣和重视。图像分类不仅仅是一种技术挑战，更是许多实际应用中必不可少的一环。随着深度学习技术的快速发展，尤其是卷积神经网络（CNN）的成功应用，图像分类在精度和效率上都取得了显著的进展。本文将聚焦于花卉图像分类任务，探索和比较不同深度学习模型在该任务上的应用和性能表现。

1.1 研究背景

图像分类作为计算机视觉中的基础问题之一，旨在将输入的图像准确分类到预定义的类别中。传统的图像分类方法依赖于手工设计的特征提取算法，如 SIFT（尺度不变特征变换）和 HOG（方向梯度直方图）。然而，这些方法在复杂场景和大规模数据集上的表现逐渐受到限制。

随着深度学习技术的兴起，特别是卷积神经网络的应用，图像分类取得了革命性的进展。深度学习模型可以通过学习从原始像素数据中提取特征，并通过层层抽象和组合，实现对复杂图像的高效分类。预训练模型的引入进一步提升了图像分类的性能，使得研究者可以在特定任务上进行快速实验和迁移学习。

2. 数据描述与预处理

2.1 数据描述

本章节首先对研究使用的花卉图像数据集进行详细描述。数据描述不仅包括数据集的来源和组成，还包括数据集的特点、图像分辨率、类别分布等方面的信息，为后续的机器学习模型训练和评估提供基础。

2.2 数据集的来源和组成

本研究使用的花卉图像数据集来源于和鲸社区。该数据集包含了 5 类花卉图像，总计 4322 张。

2.3 数据集的特点

多样性和复杂性：涵盖了多种不同种类的花卉，chamomile（雏菊）、tulip（郁金香）、rose（玫瑰）、sunflower（向日葵）、dandelion（蒲公英）。

图像分辨率：照片分辨率不高，约为 320x240 像素。图片不是缩放到一个尺寸，有不同的比例。

类别分布：每类约有 800 张照片，部分类别样本会多一些。如图 2-1。

加载的图像总数：4322
已加载标签总数：4322
{'daisy': 768, 'dandelion': 1052, 'rose': 784, 'sunflower': 734, 'tulip': 984}

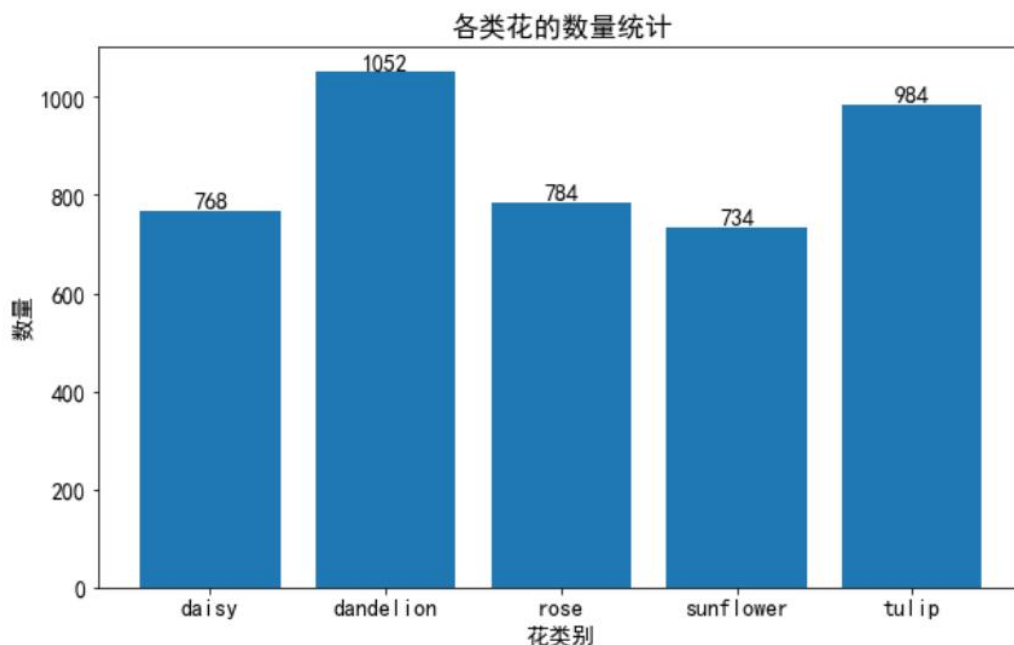


图 2-1 原数据各图片的数量统计

2.4 数据预处理

2.4.1 去重处理

为了避免训练数据中存在重复的图像样本，我们采用了基于图像哈希值的去重方法。对每张图像计算其哈希值，并将其与已有图像的哈希值进行比较。如果发现重复的哈希值，则将该图像视为重复，并从数据集中移除。

在原始加载的图像数量为 4322，标签数量也为 4322 的情况下，经过去重处理，剩余的独特图像数量为 4316，标签数量也为 4316。

2.4.2 模糊图像处理

模糊图像可能会对图像分类任务的模型训练和性能产生负面影响。因此，我们使用图像的方差作为清晰度指标。设定一个阈值，只保留那些方差高于阈值的图像，将模糊图像从数据集中过滤掉。

经过模糊图像处理，剩余的清晰图像数量为 4287，标签数量也为 4287。这一步骤确保了训练数据的质量和可用性，有助于提升模型的性能和泛化能力。

2.4.3 图像尺寸调整

为了确保所有输入图像具有相同的尺寸，我们将所有图像调整为 (224, 224) 的大小。这一步骤在模糊图像处理之前完成，以便在评估图像清晰度时保持一致性。

数据集清洗前后的统计信息总结如下：

原始加载数据： 图像数量：4322，标签数量：4322

去重处理后： 剩余独特图像数量：4316，剩余标签数量：4316

模糊图像处理： 剩余清晰图像数量：4287，剩余标签数量：4287

3.模型建立与求解

3.1 数据准备

为了进行模型的训练和评估，我们从预处理和清洗过花卉图像数据中划分出训练集和测试集。训练集用于模型参数的学习，测试集则用于评估模型的泛化能力和性能。其中 80% 用于训练，20% 用于评估。

3.2 支持向量机 (SVM)

支持向量机 (SVM) 是一种经典的监督学习模型，适用于分类和回归任务。在本研究中，我们使用线性 SVM 来处理花卉图像分类任务。

3.2.1 模型假设

在构建 SVM 模型时，我们假设每个花卉图像能够通过特征向量表示，并且这些特征向量是由 $224 \times 224 \times 3$ 维的图像数据经过预处理和展平得到的。SVM 的目标是在特征空间中找到一个最优超平面，以最大化不同类别花卉图像之间的间隔，从而实现有效的分类。

3.2.2 损失函数和优化器

我们选择交叉熵损失作为模型的损失函数，这是在多分类问题中常用的损失函数，能够衡量模型输出与真实标签之间的差异。优化器部分，我们采用随机梯度下降 (SGD) 来优化模型参数，通过调整学习率等超参数来控制模型的学习速度和稳定性。

3.2.3 模型训练

训练阶段，我们通过多次迭代训练模型，每轮迭代更新模型参数并计算训练损失和准确率。利用训练集数据进行反向传播更新模型参数，不断优化模型以提高其在训练数据上的表现。

3.3 多层感知器（MLP）

多层感知器（MLP）是一种经典的前馈神经网络模型，由多个全连接层组成。本节将详细介绍 MLP 模型的建立与求解过程。

3.3.1 模型假设

在构建 MLP 模型时，我们假设每个花卉图像通过将其像素数据展平成长度为 $224 \times 224 \times 3$ 的向量，作为模型的输入特征。MLP 模型由多个全连接层组成，通过非线性激活函数（如 ReLU）将数据映射到更高维的空间，从而实现对复杂特征的学习和提取。

3.3.2 模型定义

在本研究中，我们定义了一个包含一个隐藏层的 MLP 模型。隐藏层的大小（hidden_size）设定为 256，这是一种经验性选择，能够在保持模型复杂度适中的情况下提供足够的特征学习能力。模型的输出层为与类别数量相匹配的全连接层，用于输出每个类别的预测概率。

3.3.3 损失函数和优化器

我们选用交叉熵损失（CrossEntropyLoss）作为 MLP 模型的损失函数，用于衡量模型输出与真实标签之间的差异。优化器部分，我们选择 Adam 优化器来更新模型的参数，通过设置较小的学习率（lr=0.0001）来平衡模型的收敛速度和稳定性。

3.3.4 模型训练

训练阶段，我们使用多次迭代（epoch）来训练模型。在每个迭代中，我们通过遍历训练集数据，计算损失并进行反向传播，以调整模型参数。同时，我们实时监测并记录训练损失和准确率，以评估模型在训练过程中的表现。

3.4 残差网络（ResNet）

残差网络（ResNet）深度卷积神经网络一种经典的深度学习架构，在图像分类任务中展现了出色的性能和泛化能力。本节将详细介绍在花卉图像分类任务中应用的 ResNet 模型的建立与求解过程。

3.4.1 模型假设

我们基于以下假设构建 ResNet 模型：深度神经网络通过增加层数可以学习到更加复杂和抽象的特征表示，有助于提升图像分类任务的准确率和泛化能力。ResNet 引入了残差块(residual block)，通过跨层的直接连接(shortcut connection)来解决深层网络训练中的梯度消失问题，使得可以训练更深的网络结构。

3.4.2 模型定义

在模型建立阶段，我们选择使用预训练的 ResNet 模型作为基础架构，并针对花卉分类任务进行微调。ResNet 模型包含多个残差块，每个残差块由卷积层、批量归一化层和恒等映射或投影映射组成，最后通过全局平均池化和全连接层进行分类。

3.4.3 模型微调与训练

在模型微调阶段，我们冻结 ResNet 模型的大部分层，仅调整最后全连接层的参数，以适应花卉图像分类任务。通过多次迭代的训练，我们不断优化模型的参数，并监测训练过程中的损失和准确率变化。

3.5 模型评估

在训练阶段，通过多次迭代优化模型参数，监测损失和准确率变化。评估阶段切换为评估模式，使用测试集验证模型性能，计算准确率和 F1 分数，评估分类和泛化能力。F1 分数综合了模型的精确度和召回率，对于不均衡的类别分布是一个有用的评估指标。

4. 结果展示

4.1 SVM 模型评价

训练准确率从最初的 0.3186 增长到最终的 0.5852。模型在训练集上的预测准确性显著提升，说明其学习到了数据中的模式和特征。SVM 模型在训练过程中展现出较高的训练损失和较低的训练准确率。这可能暗示了模型对于复杂特征的学习能力较弱，或者数据特征空间的维度较高，导致难以找到一个有效的超平面来分割不同类别的花卉图像。

模型在测试集上的 F1 分数为 0.2872。表明模型在面对未见过的数据时，难以准确地进行分类，存在一定程度的欠拟合。如图 4-1。

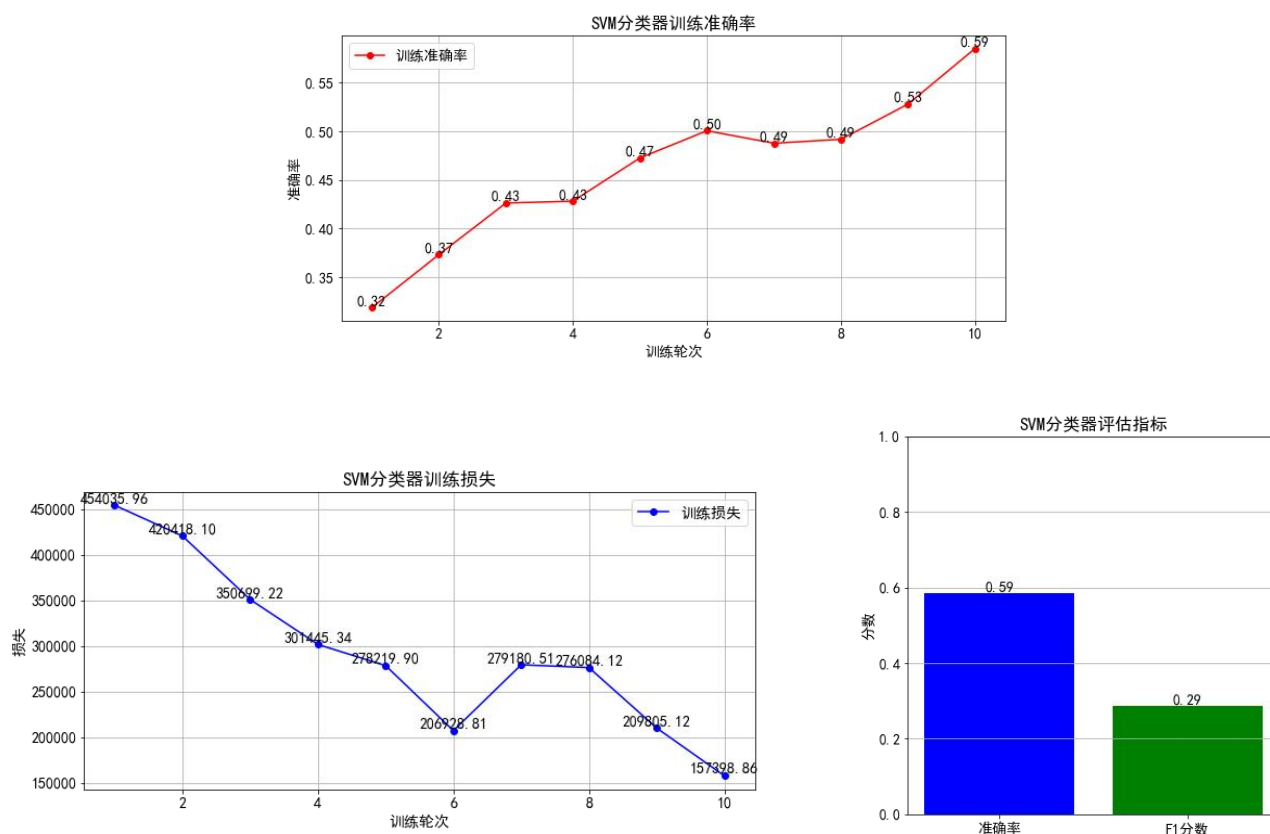
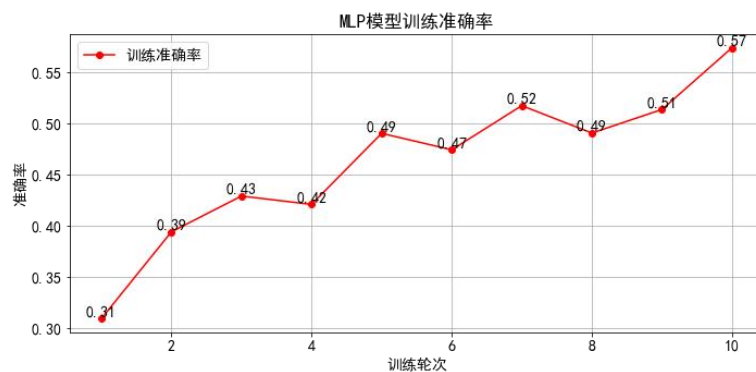


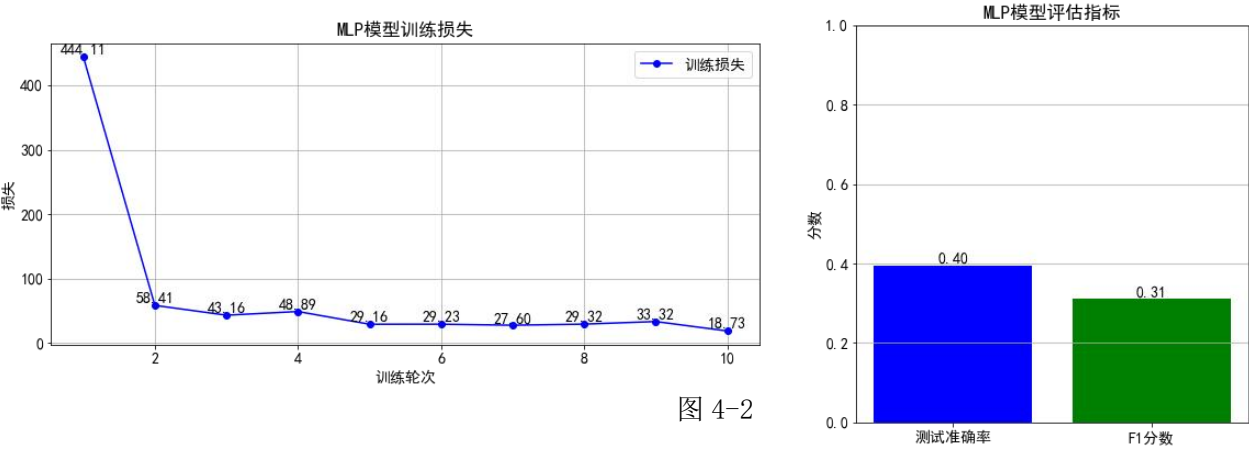
图 4-1

4.2 MLP 模型评价

MLP 模型在训练过程中表现出逐步降低的训练损失和逐步提高的训练准确率。这表明模型能够有效地学习到训练数据中的模式和特征。

模型在测试集上的准确率为 0.3956, F1 分数为 0.3112。这些指标显示出 MLP 模型在测试数据上的性能相对较弱,可能存在过拟合或者需要更多的正则化措施来提高泛化能力。如图 4-2。

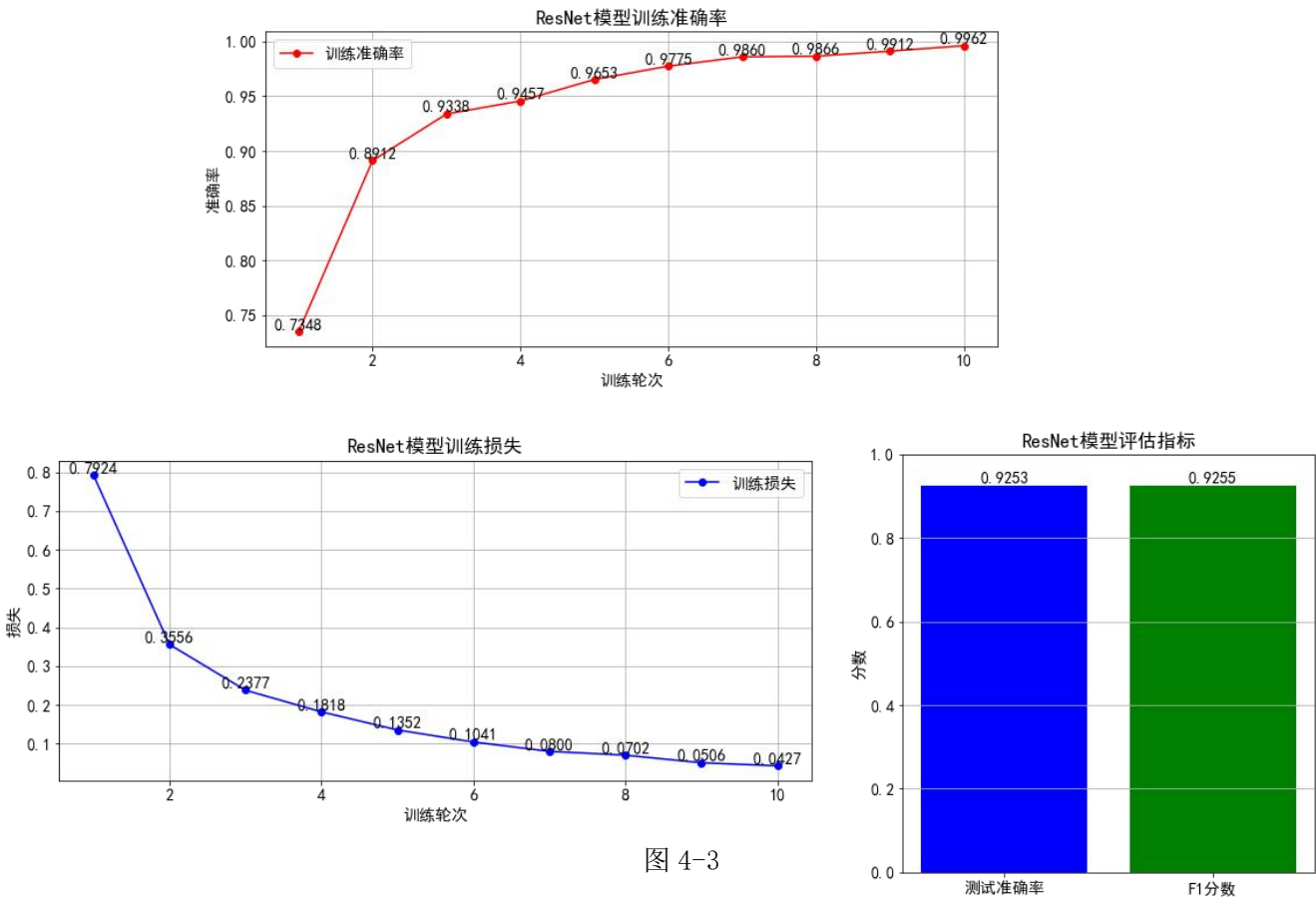




4.3 ResNet 模型评价

ResNet 模型在训练过程中表现出稳定的训练损失和高准确率的训练结果。模型逐步降低训练损失，提升训练准确率，表明其有效地学习和优化了模型参数。

模型在测试集上表现出色，准确率达到 0.9253，F1 分数为 0.9255。这显示出 ResNet 模型在分类未见过的数据时表现出了优异的泛化能力和高度的准确性。如图 4-3。



5.结论

我们评估了三种不同的机器学习模型（SVM、MLP 和 ResNet）在花卉图像分类任务中的性能。通过比较它们的训练准确率、测试准确率和 F1 分数，我们可以得出如下结论：

SVM 模型：尽管 SVM 模型在训练过程中表现出了一定的准确率提升，但在测试集上的表现较为一般。其测试准确率为 58.52%，F1 分数为 28.72%，显示出在复杂的图像分类任务中可能存在泛化能力不足的问题。

MLP 模型：MLP 模型在训练过程中训练准确率为 57.62%，但在测试集上的准确率和 F1 分数均较低，分别为 39.56%和 31.12%。这表明 MLP 模型可能在处理复杂图像特征时存在一定的过拟合问题。

ResNet 模型：相比之下，ResNet 模型在训练和测试阶段都表现出了非常优异的结果。它在训练阶段迅速收敛，并达到了 99.62%的训练准确率。在测试阶段，ResNet 模型展现出了 92.53%的测试准确率和 92.55%的 F1 分数，显示出其在花卉图像分类任务中具有较好的泛化能力和分类性能。

缺点：训练较深的 ResNet 模型可能需要更多的计算资源和时间。在此次训练的时候 SVM 和 MLP 模型基本都可以在几分钟内完成 10 轮训练，而 ResNet 模型则花费了 30 多分钟时间。可见，ResNet 模型需要足够大规模的数据集和合适的预训练模型来达到最佳效果。

ResNet 模型在花卉图像分类任务中表现最为出色，其能够有效地学习和提取复杂的图像特征，具备较强的分类能力和泛化能力。ResNet 适用于深度网络结构和复杂特征的学习，能够处理大规模数据集和复杂分类任务。残差连接有效地帮助解决梯度消失问题，使得可以训练更深的网络。因此，建议在类似的图像分类问题中考虑采用深度卷积神经网络模型，如 ResNet，以取得更好的分类效果和性能。