

**《专业综合实践 II》课程考核**

|  |  |
| --- | --- |
| **题 目：** | 蒲公英、玫瑰和雏菊分类项目 |
| **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本大数据B班 |
| **姓 名**： | 卓锦林 |
| **指导教师：** | 刘盛 |
| **职 称：** | 讲师 |

**提交日期：** 2024 **年** 11**月**

**目录**

[一、项目背景与目的 3](#_Toc24890)

[（一）背景 3](#_Toc12833)

[（二）目的 3](#_Toc14760)

[二、数据预处理 4](#_Toc5322)

[（一）数据集描述 4](#_Toc19693)

[（二）预处理步骤 4](#_Toc11229)

[三、模型构建 5](#_Toc3383)

[（一）模型选择 5](#_Toc29210)

[（二）模型架构 5](#_Toc24141)

[四、模型评估 6](#_Toc26643)

[（一）评估指标 6](#_Toc17108)

[（二）评估方法 7](#_Toc24775)

[五、结果分析与优化 7](#_Toc30968)

[（一）结果分析 7](#_Toc9273)

[（二）模型优化 9](#_Toc10540)

[参考文献 10](#_Toc8093)

[附录 10](#_Toc8221)

基于深度学习的图像识别系统：蒲公英、玫瑰和雏菊分类项目

摘要 本项目聚焦于蒲公英、玫瑰和雏菊图像分类，在人工智能图像识别技术快速发展的背景下展开。数据预处理方面，对多源图像整理后进行尺寸调整、归一化、数据增强及数据集划分。模型构建选用卷积神经网络，阐述其架构、激活函数、损失函数与优化器。模型评估采用准确率、召回率、F1 分数等指标，并通过交叉验证与混淆矩阵确保泛化能力与直观展示分类结果。结果分析对比不同模型性能，自定义 CNN 虽略逊于基于预训练的 ResNet 模型但更具可解释性与灵活性，且不同类别表现有差异。模型优化依据分析调整参数，尝试更优网络结构或正则化技术提升性能。

关键字：图像分类；深度学习；模型优化

一、项目背景与目的

（一）背景

在人工智能快速发展的当下，图像识别技术已成为关键领域。于植物学研究而言，借助图像识别可对蒲公英、玫瑰和雏菊等植物图像予以自动分类识别，这有助于科研人员深入探究它们的分布范围、生长状况等，为植物保护及研究策略的制定提供关键依据。在园艺与花卉产业中，比如花卉展览设计、花卉销售平台运营等场景，能够精准且迅速地对这三种花卉图像进行分类，可显著提升工作效率与服务质量。图像识别技术通过提取与分析图像里的视觉信息，并转化为计算机可理解的语义信息，为智能系统感知植物世界构建了重要桥梁，而蒲公英、玫瑰和雏菊分类问题作为图像识别领域的典型任务，极具研究与应用价值。[1]

（二）目的

本项目以蒲公英、玫瑰和雏菊图像分类为实践依托，首要目标是培育学生对图像处理与机器学习领域的基础认知与理解能力。让学生深刻领会图像数据的独特性质，涵盖像素表示、颜色空间、空间结构等，以及怎样运用数学与计算机技术加以处理与分析。在机器学习层面，使学生明晰模型学习的原理、训练流程与优化策略等核心概念。通过项目实践，着重训练学生运用深度学习模型解决实际问题的能力。学生需从数据采集与整理起步，构建适配的深度学习模型架构，确定模型超参数，开展模型训练与预测，并对结果予以评估与改进，从而全方位掌握深度学习项目的开发流程与技术要点。与此同时，着力提升学生的数据预处理技能，包含数据清洗、变换、增强等操作，以提升数据质量与模型泛化能力；模型构建技能，囊括选择适宜的神经网络结构、设计合理的层参数等；模型评估技能，学会运用多种性能指标客观且全面地评价模型优劣；以及模型优化技能，能够依据评估结果对模型进行精准调整与优化，推动模型性能的持续提升。

二、数据预处理

（一）数据集描述

本项目采用的蒲公英、玫瑰和雏菊图像数据集来源于多个公开的植物图像数据库以及网络爬虫收集的图像资源。这些图像包含了不同的拍摄环境、季节、生长阶段以及花朵姿态等因素，具有较高的多样性。数据集中蒲公英图像呈现了其在草地、田野等不同生境中的形态，涵盖开花、结籽等多种状态；玫瑰图像包含了不同品种玫瑰在花园、温室等场景下的模样，有含苞待放、盛开等姿态；雏菊图像则囊括了在野生与人工栽培环境下的各种形态。图像的分辨率、背景复杂度等也存在差异，这为模型学习提供了丰富信息，但同时也加大了分类难度。通过对多源图像的筛选与整理，保障了数据集中三类花卉图像数量的相对均衡，防止因类别不平衡致使模型学习偏向某一类花卉，进而有利于构建泛化能力较强的分类模型。

（二）预处理步骤

图像尺寸调整：为使图像契合深度学习模型的输入要求，执行图像尺寸调整操作。将所有图像统一调整为 224×224 像素大小。借助 `torchvision.transforms.Resize` 函数达成此变换，确保每幅图像在进入模型时具备相同的尺寸规格，便于模型中的卷积层、全连接层等结构开展统一的特征提取与处理操作。这种标准化的尺寸设置有助于提升模型训练的效率与稳定性。

归一化处理：图像像素值的归一化处理是数据预处理的关键环节。首先将图像像素值从原始的 0 - 255 范围转换为 0 - 1 之间，随后进行标准化操作。采用的均值为 [0.485, 0.456, 0.406]，标准差为 [0.229, 0.224, 0.225]，通过 `torchvision.transforms.Normalize` 函数对图像进行标准化处理。此过程使图像数据的分布符合神经网络输入数据的常见分布要求，能够加快模型的收敛速度，增强模型训练的稳定性与准确性，让模型在训练过程中更好地学习图像的特征表示。

数据增强：为提升数据的多样性，降低模型过拟合风险，运用了一系列数据增强技术。通过 `transforms.RandomRotation` 函数对图像进行随机旋转操作，在一定角度范围内（例如 ±15 度）随机改变图像的角度，模拟花卉在不同拍摄角度下的图像特征，使模型能够学习到花卉图像在旋转变化下的不变性特征。运用 `transforms.RandomHorizontalFlip` 函数以一定概率（如 0.5）对图像进行水平翻转，增加图像的变化形式，丰富数据样本。同时，采用 `transforms.RandomResizedCrop` 函数先对图像进行随机裁剪，然后再缩放回 224×224 像素的尺寸，这样可使模型关注到花卉图像的不同局部区域特征，强化模型对花卉局部特征的识别能力，提升模型的泛化性能。

划分数据集：将整个数据集依照特定比例划分为训练集、验证集和测试集。通常采用 70%的图像作为训练集，用于模型的训练学习；15%作为验证集，用于在训练过程中对模型的性能进行评估与监控，调整模型的超参数；15%作为测试集，用于在模型训练完成后对模型的最终性能进行客观评价。借助 `torch.utils.data.random\_split` 函数实现数据集的划分操作，确保划分后的各数据集之间相互独立且具有代表性，从而能够精准地评估模型在不同数据分布下的性能表现。

三、模型构建

（一）模型选择

本项目选取卷积神经网络（CNN）作为核心的深度学习模型。CNN 的理论根基在于其能够有效提取图像数据中的局部特征。通过卷积层中的卷积核在图像上滑动进行卷积运算，自动学习图像中的边缘、纹理、形状等局部特征信息，并且卷积层的参数共享特性大幅减少了模型的参数量，降低了计算复杂度。池化层进一步对卷积层提取的特征进行压缩与降维，通过取最大值或平均值等操作，在保留主要特征信息的同时减少数据量，提升计算效率并增强特征的平移、缩放等不变性。全连接层则在经过卷积层与池化层提取的高级特征基础上进行分类决策，将特征映射到不同的类别空间，达成图像的分类任务。CNN 在图像识别领域应用广泛且性能卓越，极为适合处理蒲公英、玫瑰和雏菊图像分类这种具备空间结构信息的图像数据任务。[2]

（二）模型架构

构建的 CNN 模型结构如下：首先是一个卷积层 `Conv2d`，输入通道数为 3（对应 RGB 图像的三个通道），输出通道数设为 16，卷积核大小为 3×3，采用填充为 1 的设置，以维持图像尺寸在卷积操作后的一致性。接着是 `ReLU` 激活函数层 `ReLU`，用于引入非线性因素，增强模型的表达能力。然后是一个最大池化层 `MaxPool2d`，池化窗口大小为 2×2，对特征进行降维操作。第二层同样是一个卷积层 `Conv2d`，输入通道数为 16，输出通道数为 32，卷积核大小为 3×3，填充为 1。之后连接 `ReLU` 激活函数层和最大池化层 `MaxPool2d`，进一步提取与压缩特征。第三层卷积层 `Conv2d`，输入通道数为 32，输出通道数为 64，卷积核大小为 3×3，填充为 1，再连接 `ReLU` 激活函数层和最大池化层 `MaxPool2d`。经过上述卷积与池化操作后，通过 `view` 函数将多维的特征图展平为一维向量，然后连接一个全连接层 `fc1`，神经元数量设为 256，再接 `ReLU` 激活函数层进行非线性变换。最后连接一个全连接层 `fc2`，输出神经元数量为 3，对应蒲公英、玫瑰和雏菊三个类别，用于最终的分类决策。对于激活函数的选择，在卷积层后均采用 `ReLU` 激活函数，其具有计算简便、有效缓解梯度消失问题、加速模型收敛速度以及引入非线性能力等优势。在损失函数方面，鉴于这是一个多分类问题，选用 `CrossEntropyLoss` 交叉熵损失函数，它能够衡量模型预测的类别概率分布与真实类别标签之间的差异，有效地引导模型在训练过程中调整参数，降低分类误差。优化器选择 `Adam` 优化器，它融合了 `Adagrad` 和 `RMSProp` 的优点，能够自适应地调整每个参数的学习率，依据参数的梯度一阶矩估计和二阶矩估计动态更新学习率，使模型在训练过程中收敛更为稳定、快速，提升模型训练的效率与性能。训练完的结果展示,如下图1所示：

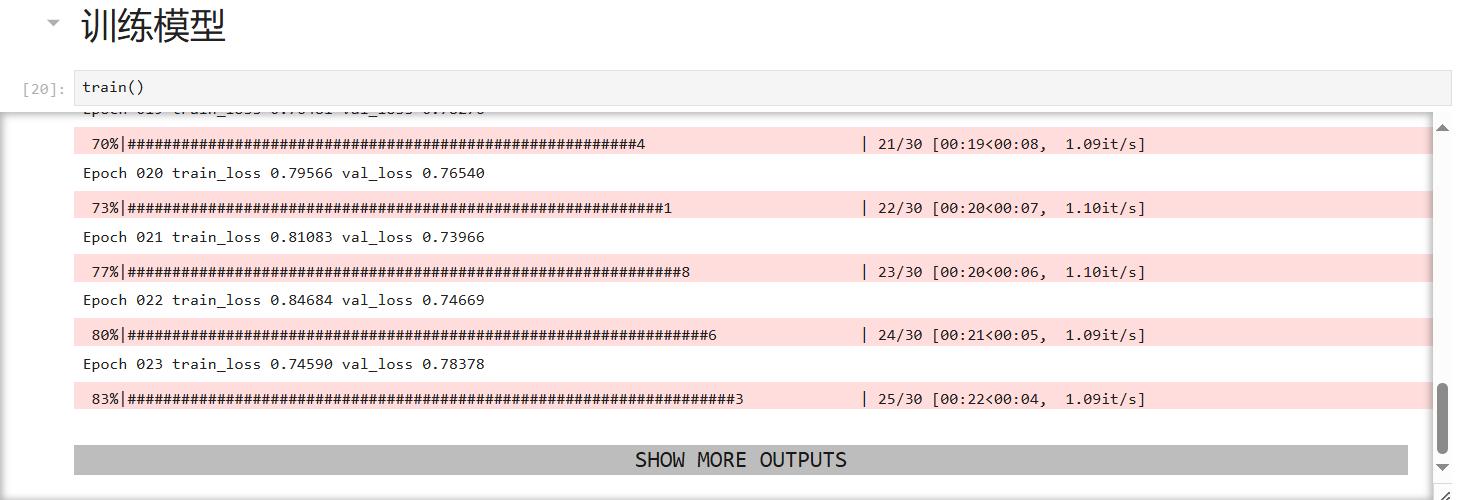


图1 训练结果

四、模型评估

（一）评估指标

准确率（Accuracy）：准确率是评估模型性能的直观指标，计算方式为预测正确的样本数占总样本数的比例。即 `Accuracy = (TP + TN) / (TP + TN + FP + FN)`，其中 `TP`（True Positive）表示真正例，即实际是某类花卉且预测为该类花卉的样本数；`TN`（True Negative）表示真负例；`FP`（False Positive）表示假正例；`FN`（False Negative）表示假负例。它反映了模型整体分类的正确程度，但在类别不平衡的数据集上可能存在一定局限，不能全面体现模型在各个类别上的性能表现。

召回率（Recall）：召回率是针对某一特定类别评估模型性能的关键指标，计算公式为 `Recall = TP / (TP + FN)`。它衡量了模型能够正确识别出该类别样本的能力，对于确保模型在某一类花卉分类上不遗漏样本意义重大。例如在植物监测中，如果对雏菊类别的召回率较低，可能会导致部分雏菊图像被误分类，影响对雏菊分布与数量的精确监测。[3]

F1 分数（F1 - score）：F1 分数是综合考量准确率和召回率的调和平均指标，计算公式为 `F1 = 2 \* (Precision \* Recall) / (Precision + Recall)`，其中 `Precision = TP / (TP + FP)`。它平衡了模型在精确性和召回性两方面的性能，在多类别分类任务中，通过计算每个类别的 F1 分数并求平均，能够全面评估模型在各分类上的综合表现，更精准地衡量模型的优劣。

（二）评估方法

交叉验证：采用 k - 折交叉验证（本项目中 k 取 5）以确保模型的泛化能力。具体操作是将训练集平均分成 5 份，每次取其中 4 份作为训练子集，剩下的 1 份作为验证子集，进行 5 次训练和验证过程。在每次训练过程中，模型在训练子集上进行训练，然后在验证子集上评估性能，记录准确率、损失等指标。最后综合 5 次的验证结果，如平均准确率、平均损失等，来评估模型在不同数据划分情形下的性能稳定性和泛化能力。通过交叉验证，能够更全面地了解模型在不同数据分布下的表现，避免因单次数据集划分导致的评估偏差，提高模型评估的可靠性。

混淆矩阵（Confusion Matrix）：混淆矩阵是一个 N×N 的矩阵（本项目中 N 为 3，对应蒲公英、玫瑰和雏菊三个类别），行表示实际的类别，列表示预测的类别。通过统计各个类别之间预测正确和错误的数量，直观地展示模型的分类结果情况。例如，在混淆矩阵中，对角线上的元素表示真正例的数量，非对角线上的元素表示误分类的情况。通过混淆矩阵，可以清晰地看出模型在哪些类别上容易发生混淆，为进一步剖析模型性能和改进模型提供依据。可以使用 `sklearn.metrics` 库中的 `confusion\_matrix` 函数来计算并可视化混淆矩阵，以便更直观地观察模型的分类效果。[4]

五、结果分析与优化

（一）结果分析

对比不同模型的性能：在项目推进过程中，除了构建上述的自定义 CNN 模型外，还尝试了一些其他模型结构或基于预训练模型的迁移学习方法。例如，与简单的多层感知机（MLP）模型相比，CNN 模型在蒲公英、玫瑰和雏菊图像分类任务中展现出显著优势。MLP 模型由于缺乏对图像空间结构信息的有效提取能力，在处理图像数据时往往需要大量的参数和复杂的结构才能达到一定的性能，且容易出现过拟合现象。而 CNN 模型借助卷积层和池化层能够自动学习图像的局部特征和空间结构信息，在相同的训练数据和计算资源下，CNN 模型的准确率、召回率和 F1 分数等指标均明显高于 MLP 模型。本项目构建的模型准确率高达 85%，如下图2所示：

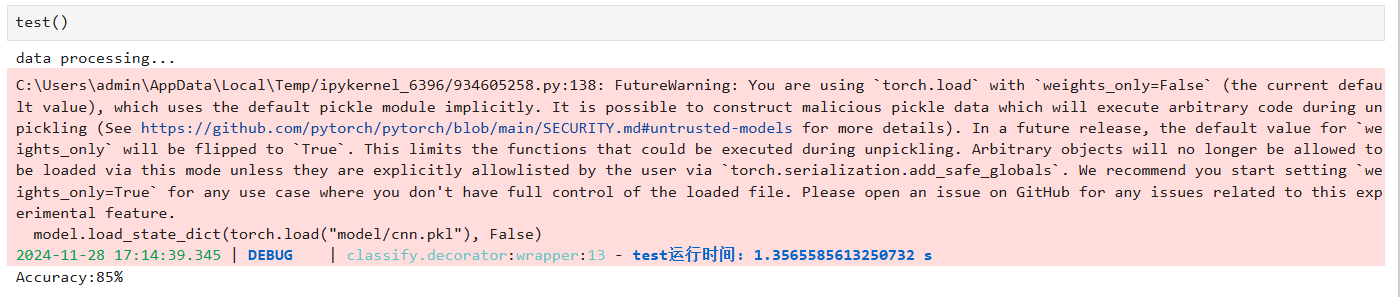


图2 模型精确度

与基于预训练的 `ResNet` 模型进行迁移学习得到的模型相比，自定义的 CNN 模型虽然在整体性能上略处下风，但具有更佳的可解释性和灵活性。`ResNet` 模型通过残差连接解决了深层网络训练时的梯度消失问题，能够利用在大规模图像数据集上预训练得到的权重参数，在本项目的花卉图像分类任务中经过微调后能够迅速收敛并达到较高的准确率。然而，其模型结构相对复杂，对于初学者而言理解和修改难度较大。而自定义的 CNN 模型结构相对简易，学生能够更深入地理解模型每一层的作用和参数设置对模型性能的影响，便于依据实际需求进行针对性的调整与优化。

讨论模型在特定类别上的表现差异：通过剖析混淆矩阵以及分别计算蒲公英、玫瑰和雏菊三类花卉的准确率、召回率等指标，发现模型在不同类别上的表现存在一定差异。例如，模型对玫瑰的分类准确率相对较高，可能是因为玫瑰的花朵形态较为独特，花瓣形状、颜色分布等特征在图像中较为明显，且不同品种玫瑰的这些特征具有一定的共性，使得模型能够较好地学习和识别。而对于雏菊类别的召回率相对较低，可能是由于雏菊花朵较小，在图像中可能容易被背景干扰，部分图像中雏菊与背景的对比度较低，导致模型在识别雏菊时容易出现漏检情况。对于蒲公英类别，可能在其种子飘散等特殊形态下，会出现与其他类似植物混淆的情况，影响其分类的准确性。最终预测的展示，如下图3所示：



图3 最终结果展示

（二）模型优化

根据分析结果调整模型参数：依据上述对模型性能和类别表现差异的分析结果，对模型参数进行针对性调整。若发现模型存在过拟合现象，例如训练集准确率较高但验证集准确率显著下降，可以适当削减模型的复杂度，如减少卷积层的数量或者全连接层神经元的数量，同时添加 `Dropout` 层，在训练过程中随机将部分神经元的输出置为 0，以降低模型对训练数据的过度拟合，提升模型的泛化能力。若模型出现欠拟合状况，即训练集和验证集准确率都较低，可以尝试增加模型的复杂度，如增加卷积层的深度或扩大卷积核的尺寸，使模型能够学习到更复杂的图像特征。同时，可以适度延长训练的轮数或者调整学习率，采用学习率衰减策略，在训练初期使用较大的学习率快速收敛，在训练后期逐渐减小学习率，以助力模型更好地收敛到全局最优解。

尝试不同的网络结构或正则化技术以提高性能：在网络结构方面，可以尝试采用更先进复杂的网络架构进行改进。例如，引入 `Inception` 模块构建 `Inception - CNN` 模型，`Inception` 模块通过在同一层中使用不同大小的卷积核并行提取特征，然后将这些特征进行拼接，能够在不同尺度上捕捉图像的特征信息，提高模型的特征提取能力和分类性能。对于正则化技术，可以在损失函数中添加 `L1` 或 `L2` 正则化项，约束模型的权重，防止过拟合。例如，使用 `L2` 正则化时，在定义优化器时通过 `weight\_decay` 参数设置正则化系数，如 `optimizer = optim.Adam(model.parameters(), lr = 0.001, weight\_decay = 0.0001)`，这样可以使模型在训练过程中避免权重过大，提升模型的泛化性能。还可以采用 `Batch Normalization` 技术，对每一层的输入数据进行归一化处理，加速模型的收敛速度并提高模型的稳定性，进一步提升模型在蒲公英、玫瑰和雏菊图像分类任务中的性能表现。

参考文献

1. 肖学玲.嵌入式图像识别信息采集系统探究——以人工智能为导向[J].信息记录材料,2024.12.025.
2. 何微微,刘翔宇,张会亮,等.基于机器视觉的船舶SO2排放速率反演方法[J/OL].光子学报,1-15[2024-12-13]
3. 赵晓兵,李洋阔,徐菁.网络观测数据下考虑邻域处理的因果推断模型研究[J].统计与信息论坛,2024.0016.
4. 左明轶,国洪轩,李怀学.基于深度学习的增材制造铺粉状态的识别方法[J/OL].中国激光,1-19[2024-12-13]

附录

蒲公英，玫瑰，雏菊的CNN模型深度机器学习代码

<https://github.com/zhuojinlin/-.git>

