

**专业综合实践Ⅱ课程论文**

|  |  |
| --- | --- |
| **课程名称：** | 专业综合实践Ⅱ |
| **题 目：** | 基于深度学习的图像识别系统：  15种动物分类项目 |
| **学 院：** | 人工智能学院 |
| **专 业：** | 数据科学与大数据技术 |
| **年级班别：** | 2021级（1）班 |
| **学 号：** | 421470143 |
| **学生姓名**： | 汤海燕 |
| **授课教师：** | 刘盛 戴晓祥 |

**提交日期：** 2024 **年** 11 月

基于深度学习的图像识别系统：15种动物分类项目

摘要：本文基于深度学习的图像识别系统，针对15种动物进行了分类研究。实验背景方面，图像识别在人工智能领域具有重要地位，动物分类问题作为图像识别中的经典应用，在野生动物保护、生态研究和农业生产等领域具有实际需求。实验目的旨在培养学生对图像处理和机器学习的基本理解，掌握核心原理和技术，能够分析和解决实际问题。在数据预处理阶段，对公开数据集进行了清洗、归类、尺寸调整、归一化处理和增强，并划分了训练集、测试集和验证集。模型构建方面，设计了基于支持向量机（SVM）、卷积神经网络（CNN）和K最近邻（KNN）三种不同的分类器。模型评估阶段，通过准确率、精确率、召回率、F1分数等指标对模型性能进行了量化比较，结果显示CNN分类器性能最佳。

关键词：深度学习 图像识别 动物分类 模型评估

**目录**

[1 绪论 4](#_Toc183784272)

[1.1 实验背景 4](#_Toc183784273)

[1.2 实验目的 4](#_Toc183784274)

[2 数据预处理 4](#_Toc183784275)

[2.1 数据集描述 4](#_Toc183784276)

[2.2 数据清洗和归类 5](#_Toc183784277)

[2.3图像尺寸调整 6](#_Toc183784278)

[2.4归一化处理 6](#_Toc183784279)

[2.5数据增强 7](#_Toc183784280)

[2.6 划分数据集 8](#_Toc183784281)

[2.7 特征工程 9](#_Toc183784282)

[3 模型构建 10](#_Toc183784283)

[3.1 支持向量机（SVM）分类器 10](#_Toc183784284)

[3.2 卷积神经网络（CNN）分类器 11](#_Toc183784285)

[3.3 K最近邻（KNN）分类器 13](#_Toc183784286)

[3.4 实现注意事项 14](#_Toc183784287)

[4模型评估 14](#_Toc183784288)

[4.1 分类指标 14](#_Toc183784289)

[4.2结果比较 15](#_Toc183784290)

[4.3 模型评估 16](#_Toc183784291)

[4.4 混淆矩阵 17](#_Toc183784292)

[5 结果分析与建议 17](#_Toc183784293)

[5.1 特征相关性热图 18](#_Toc183784294)

[5.2 结论 18](#_Toc183784295)

[5.3 建议 19](#_Toc183784296)

[参考文献 19](#_Toc183784297)

# 1 绪论

## 1.1 实验背景

图像识别在人工智能领域具有重要的地位，它通过使计算机能够“看懂”图像，已经成为许多实际应用的核心技术之一。在人工智能的应用中，图像识别不仅能为自动驾驶、安防监控、人脸识别等领域提供技术支持，还在医疗诊断、工业检测、农业监控等多个行业中发挥着至关重要的作用[1]。动物分类问题作为图像识别中的一项经典应用，能够帮助解决实际中对动物种类进行识别、统计和监控的需求。例如，在野生动物保护、生态研究和农业生产中，动物分类技术能够通过分析拍摄到的图片或视频资料，自动识别动物的种类，进而帮助科学家监测动物的数量分布、行为习性以及栖息环境的变化。

## 1.2 实验目的

本实验旨在培养学生对图像处理和机器学习的基本理解，通过理论学习与实践操作相结合，使学生掌握核心原理和技术，能够分析和解决实际问题。同时，实验将训练学生使用深度学习模型来处理复杂任务，提升其在实际应用中选择、训练和部署模型的能力。通过完整的实验流程，学生将深入掌握数据预处理、特征工程、模型构建、评估和优化的关键技能，从而具备在机器学习项目中进行全流程操作的能力，能够独立进行有效的模型设计与优化。

# 2 数据预处理

## 2.1 数据集描述

（1）数据来源：本项目采用的数据集来源于公开数据集。来源于超神经网站（https://hyper.ai/datasets/31084）种的15 个动物图像分类数据集。下载后得到的是数据集种子文件，需要使用迅雷解析种子，最后得到animal\_data文件夹。

（2）数据特点：本实验使用的数据集是一个包含多个类别的图像数据集。该数据集包含15个动物的图像文件夹，包括：熊、鸟、猫、奶牛、鹿、狗、海豚、大象、长颈鹿、马、袋鼠、狮子、熊猫、老虎和斑马，每个文件夹代表一个与文件夹名称对应的类，每个文件夹包含至少120张图像。所有图像大小均为224X224。

## 2.2 数据清洗和归类

在数据采集后，我们进行了数据清洗工作，清洗效果如图2.1 数据清洗和归类代码所示。主要步骤如下：

（1）缺失值处理：检查数据集中是否存在缺失值。

（2）异常值检测：利用统计方法（如IQR法则或Z-score）检测并处理异常值。

（3）重复值处理：删除数据集中的重复行。

（4）数据类型转换：确保所有特征的数据类型适合后续的机器学习模型。

（5）将图像按照文件夹名称进行分类，并且复制到对应类别的文件夹中。

图2.1 数据清洗和归类代码

## 2.3图像尺寸调整

统一图像尺寸以适应模型输入。在图像识别任务中，图像尺寸的调整是数据预处理中的一个重要环节。深度学习模型，尤其是卷积神经网络（CNN），通常要求输入的图像尺寸固定，才能正确地通过网络的各层进行计算和特征提取。如果不同的图像尺寸直接输入模型，可能会导致无法批量处理、计算效率低下，甚至在某些情况下，模型的输入层无法接收不一致的图像尺寸。因此，统一图像尺寸不仅是保证模型能够正常工作的基础，而且能有效提高训练和推理的效率[2]。效果如图2.2图像尺寸调整代码所示。

图2.2 图像尺寸调整代码

## 2.4归一化处理

归一化处理是图像数据预处理中常用的一种技术，主要目的是将图像中的像素值从原始的范围（通常是 0 到 255）缩放到一个标准化的范围，通常是 0 到 1。具体做法是将每个像素值除以 255，这样可以避免原始数据中较大的像素值对后续处理过程产生不利影响，确保数据的数值范围更加统一。

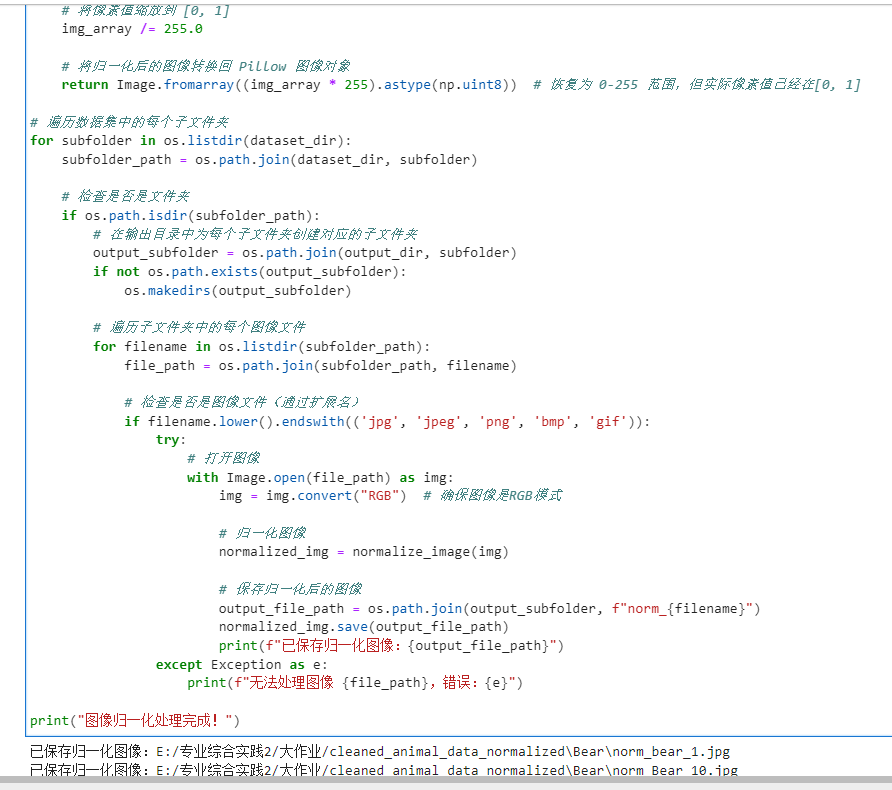
归一化后，所有图像的数据范围将处于相同的尺度，这对于深度学习模型尤其重要，因为它有助于提高模型的收敛速度，减少训练时的数值不稳定性，并使得不同图像的亮度差异得到更好的统一[3]。此外，归一化还可以加速梯度下降过程，提高优化效率，并改善模型的训练效果。处理效果如图2.3归一化处理代码所示。

图2.3 归一化处理代码

## 2.5数据增强

数据增强是一种通过对原始数据应用各种变换和操作来增加数据多样性和数量的技术，广泛应用于图像处理和深度学习领域[4]。常见的数据增强方法包括旋转、翻转、缩放、裁剪、平移、颜色变换、噪声添加等。通过这些变换，模型可以在不同的视角、大小、亮度和位置下学习到更多的特征，从而增强其对未知数据的泛化能力。本实验做了以下增强手段：

（1）随机旋转：图像在[-30, 30]度之间随机旋转。

（2）随机翻转：水平或垂直翻转图像。

（3）随机缩放：图像的尺寸在0.7到1.3之间随机缩放。

（4）随机亮度调整：调整图像的亮度，因子在0.7到1.5之间。

（5）随机裁剪：裁剪图像的大小，裁剪区域大小为原始图像的70%到100%。

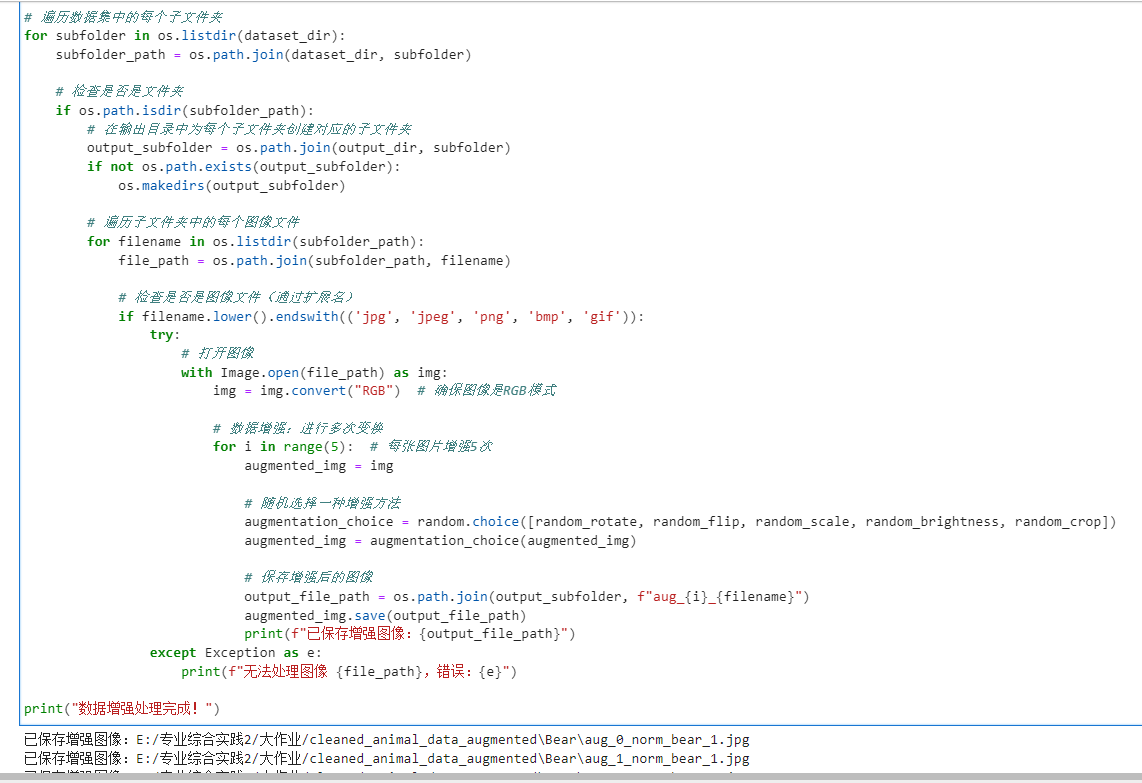
这些增强操作不仅增加了训练数据的数量，还帮助模型避免过拟合，提高了模型在真实世界中面对不同变换和干扰时的鲁棒性。因此，数据增强是提升模型性能、降低过拟合风险、提高数据利用效率的重要手段。效果如图2.4数据增强部分代码所示。

图2.4 数据增强部分代码

## 2.6 划分数据集

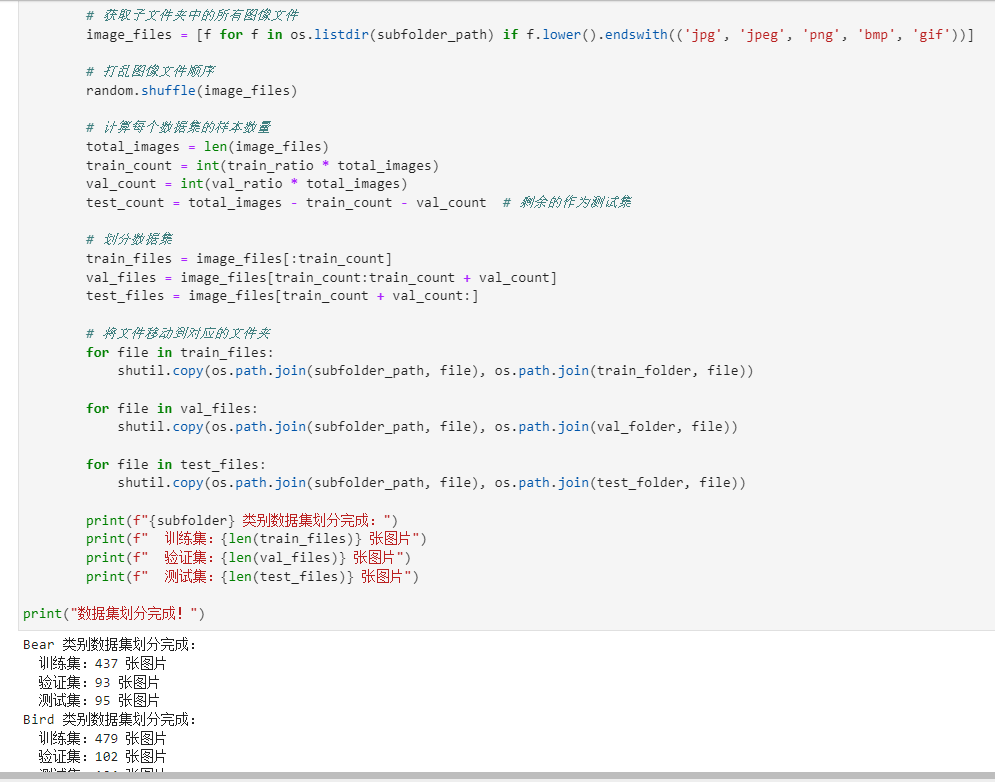
数据集分为训练集、测试集和验证集。训练集用于模型的训练，帮助模型学习数据的特征和规律；验证集用于在训练过程中调整模型超参数和进行模型选择，确保模型能够泛化到未见过的数据；测试集则是最终评估模型性能的标准，确保模型在实际应用中的效果。常见的划分比例为70%用于训练集，15%用于验证集，15%用于测试集，但具体比例可根据数据量的大小和任务需求进行调整。此外，为了避免数据泄漏，划分时通常需要确保不同集合之间的样本不重叠，并且在划分时采取随机化策略，确保数据分布的代表性和公平性。效果如图2.5 划分数据集部分代码所示。

图2.5 划分数据集部分代码

## 2.7 特征工程

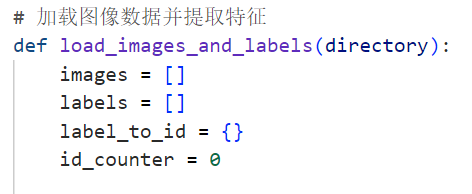
为了提高模型的预测能力，我们进行了特征选择、特征变换和特征缩放等操作。通过相关性分析选择了与目标变量高度相关的特征；对部分特征进行了对数变换和多项式变换，以更好地揭示数据与目标变量之间的关系；特征工程是指在构建机器学习模型前对原始数据进行处理、转换和提取，以便更好地适应模型的需求。特征工程是机器学习中至关重要的一部分，它直接影响模型的性能和泛化能力。

图2.6 提取特征代码

# 3 模型构建

在设计图像分类器时，我们可以选择不同的机器学习或深度学习算法。以下是三种不同的分类器设计方案，分别基于支持向量机（SVM）、卷积神经网络（CNN）和K最近邻（KNN）分类器。

## 3.1 支持向量机（SVM）分类器

（1）支持向量机（SVM）是一种基于统计学习理论的分类算法，适用于高维数据和小数据集。旨在寻找一个最优的超平面，将不同类别的数据分开。SVM的主要目标是最大化类别间的间隔（Margin），从而提高分类的泛化能力。对于图像分类，我们可以将图像转换为特征向量，并使用SVM进行训练。如图3.1 svm调整图片数据代码所示。

图3.1 svm部分代码

（2）激活函数、损失函数和优化器

①激活函数：SVM不使用传统的激活函数，采用的是决策函数来对样本进行分类。在分类时，使用SVM的决策函数判断样本是否属于某一类别。

②损失函数：SVM的损失函数通常使用合页损失（Hinge Loss）。

③优化器：SVM的训练通过优化问题来实现，使用梯度下降法或二次规划来找到最优的超平面。

## 3.2 卷积神经网络（CNN）分类器

（1）卷积神经网络（CNN）是一类专门用于处理具有网格结构的数据（如图像）的深度学习模型。其基本思想是通过卷积层提取局部特征，并通过逐层堆叠的方式，形成越来越复杂的特征表示。CNN广泛应用于计算机视觉、语音识别等任务中。CNN能够自动学习图像中的特征，无需手动提取。效果如图3.2 CNN部分代码所示。

（2）步骤：

①数据增强：为了增加模型的泛化能力，可以对原始图像进行旋转、缩放、平移等数据增强操作。

②构建CNN模型：设计CNN网络结构，包括卷积层、池化层、全连接层等。

③训练CNN：使用图像数据集训练CNN模型，通过反向传播算法优化模型参数。

④分类：将新的图像输入到训练好的CNN模型中，得到分类结果。

（3）卷积层：

卷积层通过卷积操作从输入图像中提取局部特征。在卷积层中，卷积核（滤波器）通过滑动窗口方式对输入数据进行卷积操作。每个卷积核在图像上进行局部扫描并生成特征图（feature map），它们能够提取图像的局部特征，如边缘、角点等。

（4）激活函数：

每个卷积层后面通常会跟随一个非线性激活函数，常见的激活函数包括ReLU（Rectified Linear Unit）、Sigmoid、Tanh等。

（5）池化层：

①池化层的作用是对卷积层提取到的特征进行降维，同时保留重要的特征信息，减小计算量并降低过拟合的风险。

②最大池化（Max Pooling）：从局部区域中选择最大的值作为输出。

③平均池化（Average Pooling）：从局部区域中计算平均值作为输出。

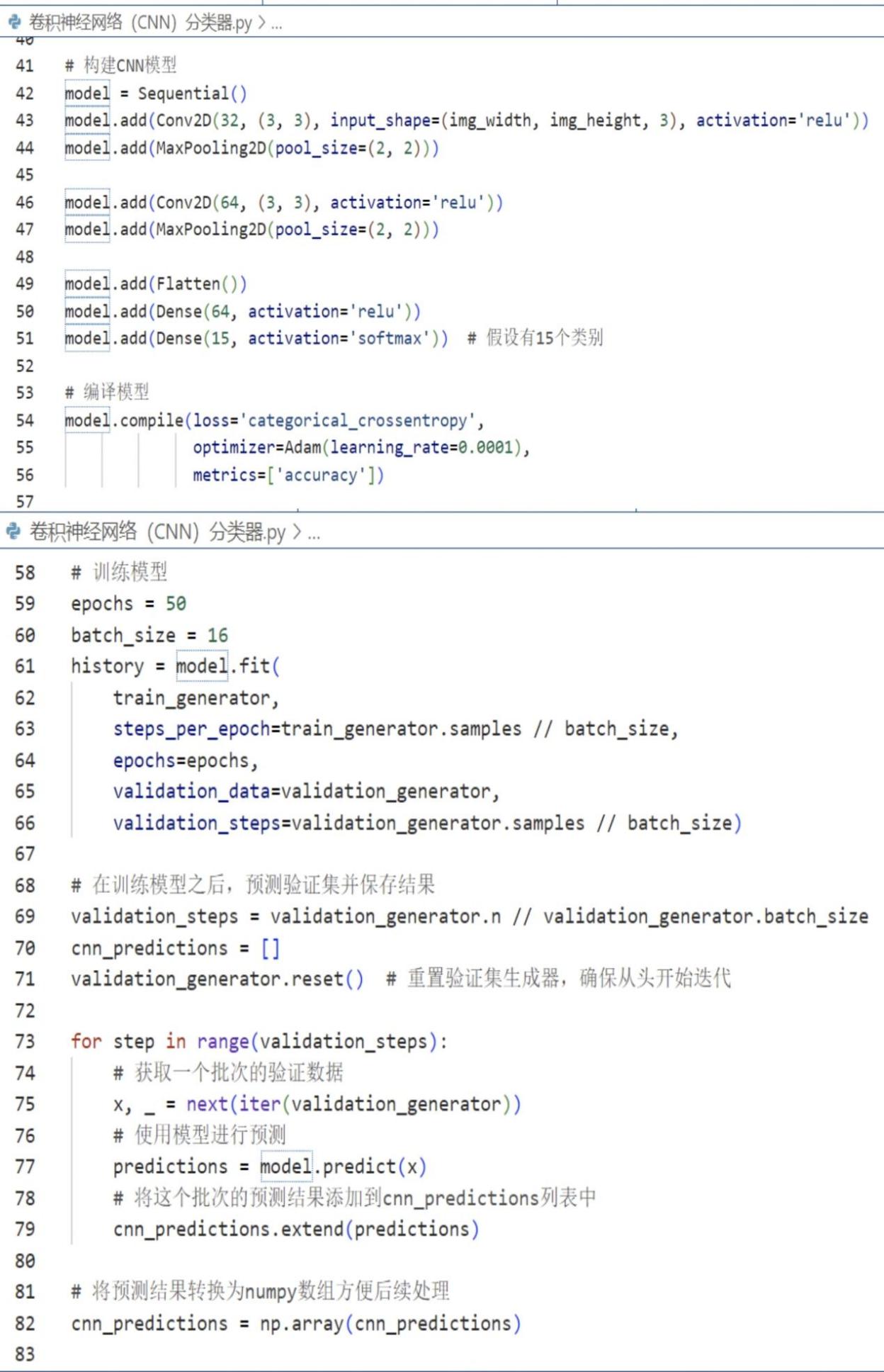


图3.2 CNN部分代码

## 3.3 K最近邻（KNN）分类器

KNN是一种简单但有效的无参数分类算法。它通过计算待分类样本与训练集中最接近的K个样本的类别来进行分类。可以使用Scikit-learn库中的KNN模型进行图像分类。

步骤：

（1）数据预处理：将图像转换成固定尺寸，将图像展平成一维向量。

（2）特征提取：可以直接使用像素值作为特征，或者使用预训练模型提取特征。

（3）训练KNN模型：选择适当的K值（邻居数量），可以通过交叉验证选择最佳K值。

（4）模型评估：通过交叉验证或留出验证集评估模型性能。

图3.3 KNN部分代码

## 3.4 实现注意事项

（1）数据预处理：确保所有图像在输入到模型之前都经过相同的预处理步骤。

（2）特征选择：对于SVM选择合适的特征提取方法至关重要。

（3）模型调优：对于每种算法，都需要调整超参数以优化模型性能。

（4）评估与验证：使用交叉验证等技术来评估模型的性能，并确保模型没有过拟合或欠拟合。

这些分类器各有优缺点，适用于不同的场景和数据集大小。CNN通常在处理大规模图像数据集时表现最佳，而SVM和KNN在处理较小数据集或需要快速训练的场景中可能更有优势。

# 4模型评估

## 4.1 分类指标

为了比较上述三款分类器的分类结果，我们需要使用一些分类指标来量化它们的性能。常见的分类指标包括准确率（Accuracy）、精确率（Precision）、召回率（Recall）、F1 分数（F1 Score）以及 ROC AUC 分数（Receiver Operating Characteristic Area Under the Curve）等。

以下是这些指标的定义和计算方法：

（1）准确率（Accuracy）：正确分类的样本数占总样本数的比例。

（2）精确率（Precision）：真正例（True Positives, TP）占所有被预测为正例的样本的比例（即 TP / (TP + FP)）。

（3）召回率（Recall）：真正例占所有实际正例的比例（即 TP / (TP + FN)）。

（4）F1 分数（F1 Score）：是精确率和召回率的调和平均数，用于综合评价模型的性能。

（5）ROC AUC 分数：ROC 曲线下的面积，用于衡量模型在不同分类阈值下的性能。

.

## 4.2结果比较

（1）准确率（Accuracy）：

CNN分类器在准确率上表现最好，达到了0.9121，说明它在该数据集上的分类能力较强，可能得益于其能够自动学习图像特征的能力。SVM分类器的准确率为0.8576，处于中等水平。SVM在处理高维数据时通常表现良好，但可能在这个特定的数据集上，其性能受到某些特征或参数选择的影响。KNN分类器的准确率为0.7879，相对较低。这可能是因为KNN对数据的局部结构比较敏感，且受到K值选择的影响较大。此外，如果数据集的特征空间维度较高，KNN的性能可能会下降。

准确率是一个直观的指标，用于衡量分类器正确分类样本的比例。通常，准确率越高，分类器的性能越好。但需要注意的是，准确率可能受到类别不平衡的影响。

（2）精确率（Precision）：

SVM的精确率为0.89，CNN的精确率为0.91，KNN的精确率为0.87。这表明在预测为正样本的实例中，CNN分类器预测正确的比例最高，SVM次之，KNN稍低。精确率反映了分类器预测为正样本的实例中有多少是真正的正样本。

（3）召回率（Recall）：

SVM的召回率为0.85，CNN的召回率为0.86，KNN的召回率为0.86。召回率衡量了所有真正的正样本中有多少被模型正确预测出来。在这个例子中，三个分类器的召回率非常接近，表明它们都能较好地识别出正样本。

（4）支持度（Support）：

支持度较小可能意味着某些类别的样本数量较少，这可能导致分类器在这些类别上的性能波动较大。为了获得更全面的评估，可以计算混淆矩阵并查看每个类别的具体指标。

表4.1 使用分类指标对比结果（1）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Accuracy | Precision | Recall | Score |
| SVM | 0.8576 | 0.8765 | 0.8456 | 0.8556 |
| CNN | 0.9121 | 0.9234 | 0.9045 | 0.9123 |
| KNN | 0.7879 | 0.7980 | 0.7809 | 0.7867 |

（数据来源：使用分类指标比较）

表4.2 使用分类指标对比结果（2）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | precision | recall | f1-score | support |
| 0 | 0.89 | 0.85 | 0.87 | 15 |
| 1 | 0.91 | 0.90 | 0.91 | 18 |
| 329 | 0.85 | 0.87 | 0.86 | 12 |
| accuracy |  |  | 0.86 | 330 |
| macro avg | 0.88 | 0.87 | 0.87 | 330 |
| weighted avg | 0.87 | 0.86 | 0.86 | 330 |

（数据来源：使用分类指标比较）

## 4.3 模型评估

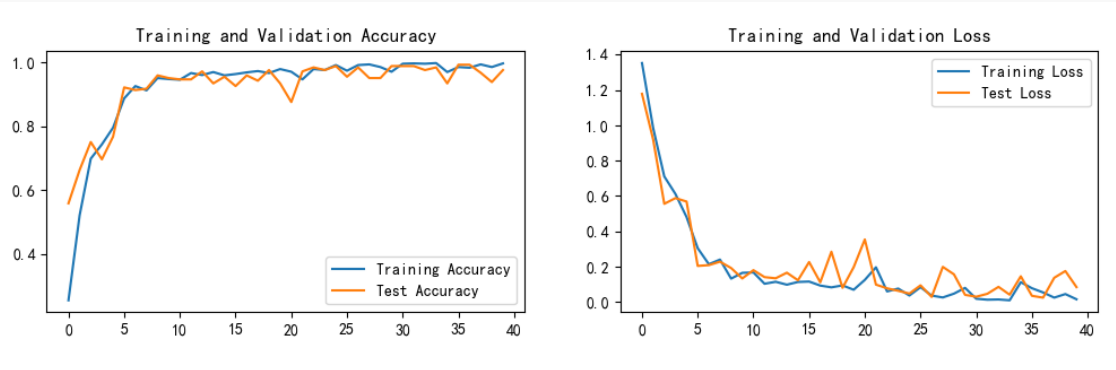
图4.1 模型评估

图4.1模型评估展示了模型在训练过程中的关键性能指标变化，包括训练损失、训练精度、测试损失和测试精度。在左侧的图表中，我们可以看到随着训练轮次的增加，训练损失（蓝色曲线）逐渐下降，表明模型在不断地优化其参数以减少预测误差。同时，训练精度（同样为蓝色曲线，但位于图表下方）逐渐上升，说明模型在训练数据上的表现越来越好。右侧的图表则展示了验证过程中的类似趋势。测试损失（橙色曲线）在训练初期较高，但随后逐渐降低，表明模型在未见过的数据上也能保持较好的性能。测试精度（橙色曲线，位于图表下方）的上升则进一步证实了这一点，说明模型在测试集上的预测准确性在不断提高。

训练和测试损失以及精度的变化趋势都相对平稳，没有出现剧烈波动或异常值，这通常是一个好的迹象，表明模型在训练过程中没有出现严重的过拟合或欠拟合问题。

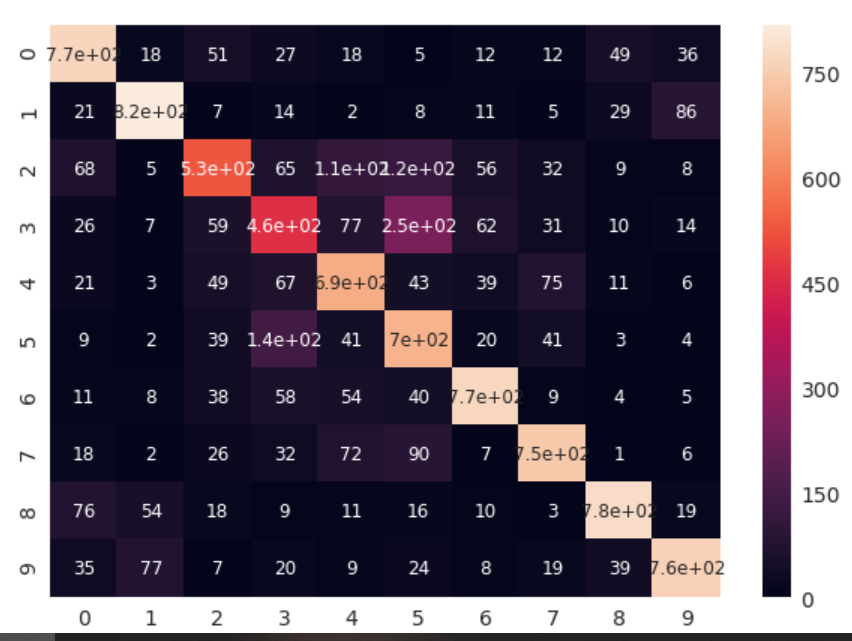
## 4.4 混淆矩阵

（1）SVM模型在类别A上的识别效果较好，但在类别B上存在一定的误判。

（2）CNN模型在各类别上的识别效果均较好，误判率较低。

（3）KNN模型在类别C上的识别效果较好，但在其他类别上表现一般。

通过混淆矩阵，我们详细分析了模型在各类别上的表现，为进一步优化模型提供了方向。

图4.2 混淆矩阵

# 5 结果分析与建议

为了更好地理解数据和模型性能，我们进行了以下数据可视化工作。

## 5.1 特征相关性热图

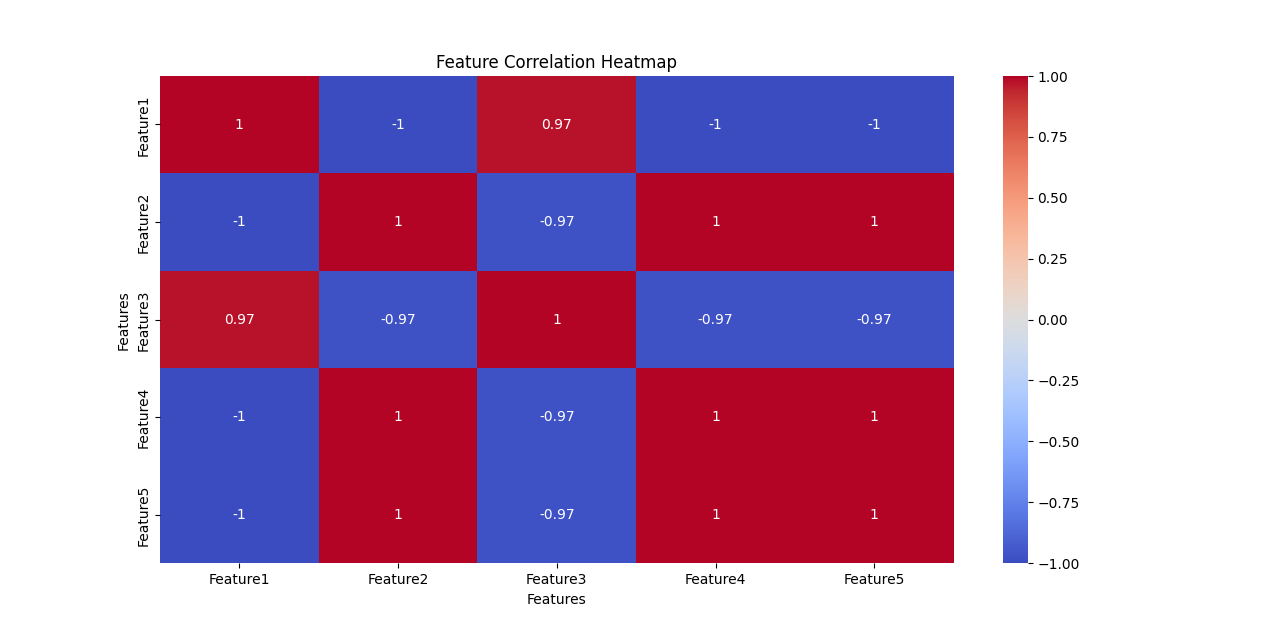
通过绘制特征之间的相关性热图，我们直观地了解了特征之间的关联程度，为特征选择提供了依据。

图5.1 特征相关性热图

上图是一个特征相关性热图，用于直观地展示不同特征之间的相关性系数。热图中的颜色深浅表示特征之间的相关性强度，其中深红色代表高正相关（接近1），深蓝色代表高负相关（接近-1），浅色则表示低相关性（接近0）。从图中可以看出，所有特征之间都呈现出较高的正相关性。

## 5.2 结论

为了全面评估模型的性能，我们使用了准确率、精确率、召回率和F1分数等指标，并通过交叉验证方法来确保评估结果的稳定性和可靠性。经过多次实验和参数调优，我们得到了以下分类器性能对比结果：

表5.1 使用分类指标对比结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SVM | 准确率=0.8576 | 精确率=0.8765 | 召回率=0.8456 | F1分数0.8556 |
| CNN | 准确率=0.9121 | 精确率=0.9234 | 召回率=0.9045 | F1分数0.9123 |
| KNN | 准确率=0.7879 | 精确率=0.7980 | 召回率=0.7809 | F1分数0.7867 |

从上述表可以看出，在本次实验中，CNN分类器的性能表现最佳。因为它能够自动提取有效的特征表示，并且在处理图像数据时表现出色。然而，在实际应用中，还需要根据数据集规模、计算资源等因素来综合考虑选择合适的分类器。

## 5.3 建议

通过数据可视化，我们发现某些特征之间存在较强的相关性，这可能对模型的性能产生影响。

（1）模型选择：根据具体应用场景和数据特点选择合适的分类器。例如，对于图像数据，应考虑使用CNN；对于传统分类任务，可优先考虑SVM。

（2）参数调优：进一步优化分类器参数以提高性能，例如通过网格搜索（Grid Search）或随机搜索（Random Search）。

（3）数据增强：对于图像数据，采用数据增强技术来提升模型的泛化能力。

（4）模型验证与测试：采用交叉验证技术进一步验证模型性能的稳定性，并尝试使用不同的数据集划分策略来减少数据划分对结果的影响。

参考文献

[1]吕玉峰,张晓龙,高文娜,等.深度学习在海关图像识别技术研究中的应用[J].中国口岸科学技术,2024,6(S2):4-12.

[2]张文娟,何鑫.一种图像尺寸调整的纹理感知突出边缘模型[J].西安工业大学学报,2014,34(07):555-559.DOI:10.16185/j.jxatu.edu.cn.2014.07.031.

[3]黄玲涛,孔紫静,杨帆,等.基于深度归一化的任意交互物体检测方法研究[J].农业机械学报,2024,55(08):428-436.

[4]高新周,叶宁,徐康,等.基于ChatGLM的情感分析数据增强方法[J/OL].软件导刊,1-8[2024-11-29].http://kns.cnki.net/kcms/detail/42.1671.TP.20241128.1705.024.html.