

**基于机器学习的垃圾图像分类**

|  |  |
| --- | --- |
| **题 目：** | 基于机器学习的垃圾图像分类 |
| **学 院：** | 数据科学学院 |
| **专 业：** | 数据科学与大数据技术 |
| **年级班别：** | 数据科学与大数据2021级（1）班 |
| **学 号：** | 421470150 |
| **学生姓名**： | 吴奇芮 |

**提交日期：** 2024 **年** 6 月

# 1 绪论

## 1.1 选题背景

在现代社会中，垃圾分类和环境保护成为了全球范围内的热点议题。随着城市化进程的加速和生活水平的提高，人们产生的垃圾数量不断增加，如何高效地对垃圾进行分类处理成为了一项迫切的任务。传统的垃圾分类方式依赖于人工，耗时耗力且易出错，因此迫切需要借助现代技术手段实现自动化垃圾分类。

近年来，随着计算机视觉和机器学习技术的快速发展，基于图像识别的垃圾分类成为了一种新的解决方案。通过将垃圾图像与已知类别进行比对，机器学习模型可以自动识别出垃圾的类型，从而实现自动化的垃圾分类。这不仅能够提高垃圾分类的效率，还有助于减少人工错误和提升环境保护意识。

本研究旨在探索基于机器学习的垃圾图像分类方法，通过构建和优化机器学习模型，实现对垃圾的准确分类，为垃圾处理和环境保护提供技术支持。

## 1.2 研究内容

本研究的主要内容包括：

（1）收集和整理垃圾图像数据集：从不同来源获取包含各类垃圾图像的数据集，并进行数据清洗和整理，为后续的模型训练做准备。

（2）图像预处理和增强：对图像进行统一大小调整、归一化处理，并进行数据增强以扩充数据集，提高模型的泛化能力。

（3）构建和优化机器学习模型：包括K近邻（KNN）、梯度提升树（XGBoost）和卷积神经网络（CNN）等模型的构建与优化，通过网格搜索等技术确定最佳参数。

（4）模型评估和比较：使用准确率、精度、召回率和F1值等指标对各模型进行评估，并比较它们在垃圾图像分类任务上的性能。

# 2 数据收集与处理

## 2.1 数据收集

### 2.1.1 数据收集方法

本研究采用手机收集数据的方法，主要包括拍摄和手机网络收集两种方式。通过手机摄像头对不同类别的垃圾进行拍摄，包括其他垃圾（一次性快餐盒）、厨余垃圾（茶叶渣）、可回收物（塑料衣架、毛绒玩具、锅）以及有害垃圾（干电池、软膏）等。拍摄过程中，尽量保证图像清晰度和光线条件的一致性，以提高后续图像处理和模型训练的效果。利用手机应用程序或通过手机浏览器访问相关网站，收集包含垃圾图像的网络数据。这些数据可能来源于公共数据库、社交媒体平台或其他网络资源，通过下载或截图的方式获取图像数据，以丰富研究数据集。



图2-1数据集示例

以上两种方式结合使用，能够获取到多样化的垃圾图像数据，为后续的数据处理和模型训练提供充足的数据支持。

### 2.1.2 数据内容描述

本研究的数据集包含7种常见生活垃圾：，包括其他垃圾（一次性快餐盒）、厨余垃圾（茶叶渣）、可回收物（塑料衣架、毛绒玩具、锅）以及有害垃圾（干电池、软膏）。

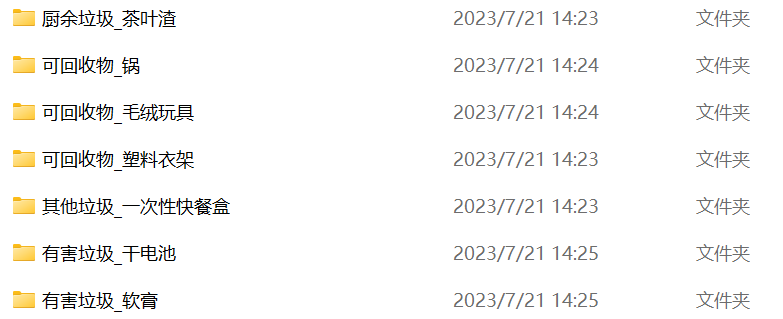


图2-2数据分类文件夹

每个类别图片的数量如下表所示。

表2-1不同类别图像数量

|  |  |
| --- | --- |
| 类别 | 图片数量 |
| 其他垃圾\_一次性快餐盒 | 232 |
| 厨余垃圾\_茶叶渣 | 377 |
| 可回收物\_塑料衣架 | 302 |
| 可回收物\_毛绒玩具 | 255 |
| 可回收物\_锅 | 385 |
| 有害垃圾\_干电池 | 312 |
| 有害垃圾\_软膏 | 381 |

## 2.2 数据处理

### 2.2.1 图像统一大小

在数据处理过程中，我们对图像进行了统一大小的处理。通过将所有图像调整为相同的尺寸（128x128像素），可以消除图像大小差异对模型训练的影响，同时也有利于加快图像处理和模型训练的速度。

### 2.2.2 图像归一化

为了进一步优化图像数据，我们对图像进行了归一化处理。归一化操作将图像像素值缩放到[0,1]范围内，有利于提高模型的稳定性和收敛速度，减少训练过程中的梯度消失或梯度爆炸问题。

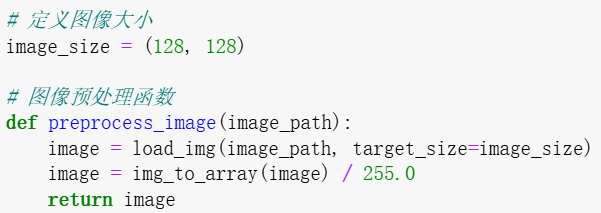


图2-3图像归一化代码

### 2.2.3 图像增强

为了扩充数据集并提高模型的泛化能力，我们采用了图像增强的技术。通过对图像进行随机旋转、平移、缩放、水平翻转等操作，生成新的图像样本，从而增加了数据集的多样性，使模型更好地适应各种场景和变化。



图2-4图像增强代码

### 2.2.4 数据集划分

最后，我们将处理好的数据集划分为训练集和测试集。通常采用的划分比例为80%的数据用于训练模型，20%的数据用于评估模型的性能。这样的划分方式既保证了模型训练的充分性，又能有效评估模型的泛化能力和准确性。

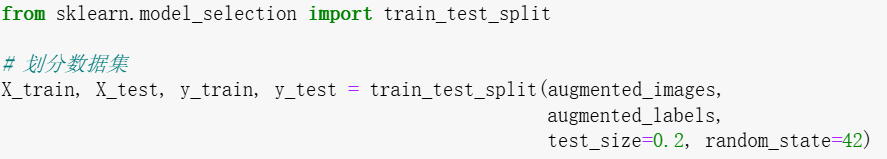


图2-5 图像数据划分代码

# 3 机器学习模型构建评估

## 3.1 KNN模型构建

### 3.1.1 KNN模型原理

K近邻（K-Nearest Neighbors，KNN）是一种基本的分类与回归方法，其原理简单直观。对于一个待分类的样本，KNN算法会计算它与训练集中各个样本的距离，并选择距离最近的K个样本作为其邻居。通过对这K个邻居样本的标签进行投票或加权投票，来决定待分类样本的类别。

KNN模型不需要训练过程，直接利用训练集的样本进行预测，因此计算复杂度较高，但模型理解和实现较为简单，适用于小规模数据集和非线性问题。

### 3.1.2 KNN模型参数调优

在构建KNN模型时，我们通过调整K值和距离度量方法来优化模型。

K值调优：我们使用网格搜索（Grid Search）方法，在K的范围内搜索最优的K值。通过交叉验证（Cross-Validation）评估每个K值下模型的性能，并绘制调优曲线。

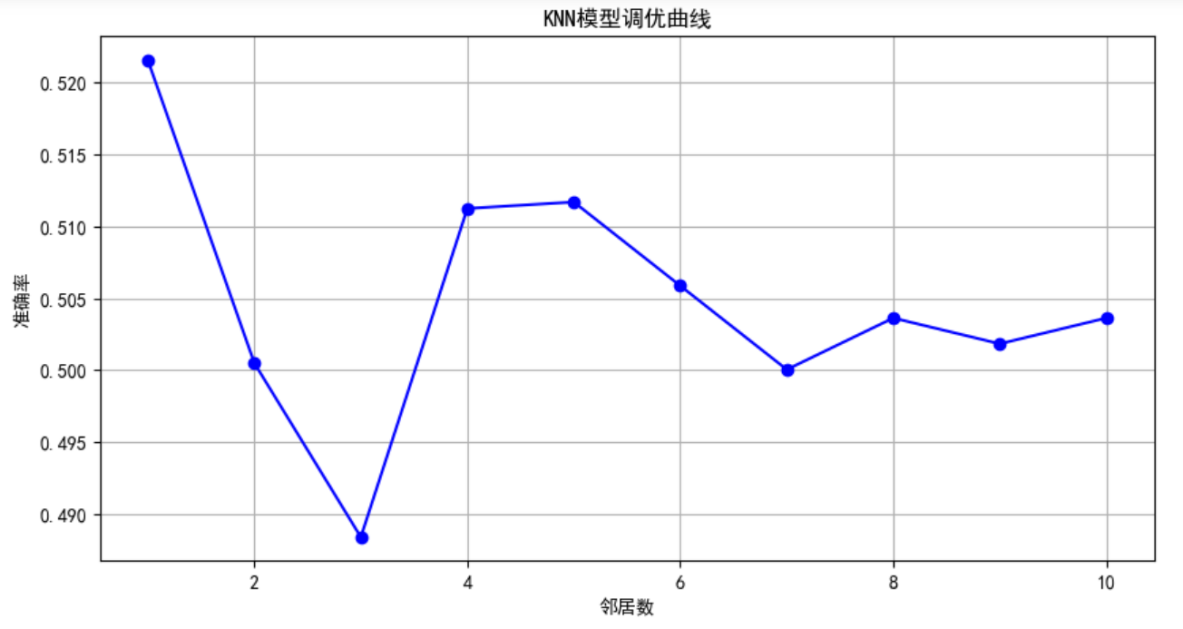


图3-1 KNN模型参数调优曲线图

调优曲线结果分析：经过调优曲线的分析，我们发现当K=1时模型性能达到最优，此时模型的准确率和泛化能力较好，因此确定K=1为最优参数。

## 3.2 XGBoost模型构建

### 3.2.1 XGBoost模型原理

梯度提升树（Gradient Boosting Decision Tree，GBDT）是一种基于决策树的集成学习算法，而XGBoost是在GBDT基础上进行了优化和改进的算法。XGBoost采用了正则化项和二阶导数信息等技术，能够更好地处理大规模数据和复杂模型，具有较高的性能和效率。

XGBoost通过不断迭代训练弱分类器（决策树），每次迭代都根据前一轮的残差拟合新的分类器，最终将多个弱分类器组合成一个强分类器，实现对复杂数据的精确预测。

### 3.2.2 XGBoost模型参数调优

在构建XGBoost模型时，主要调优最大树深度（Max Depth）：限制决策树的最大深度，防止过拟合，同时也影响模型的复杂度和泛化能力。

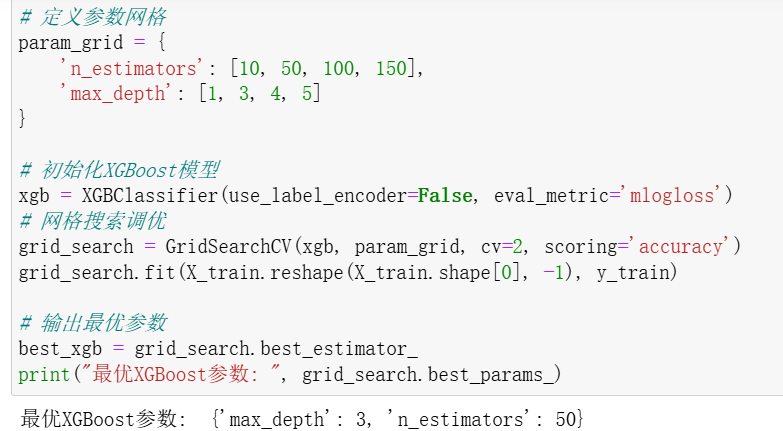


图3-2 XGBoost模型参数调优代码及结果

调优结果分析：经过参数调优，我们确定了最优参数组合，最大树深度为3，。这些参数的选择既保证了模型的性能和稳定性，又有效防止了过拟合现象，使得模型在测试集上的表现较为优秀。

## 3.3 CNN模型构建

### 3.3.1 CNN模型原理

卷积神经网络（Convolutional Neural Network，CNN）是一种专门用于处理图像数据的深度学习模型。CNN通过卷积层（Convolutional Layer）、池化层（Pooling Layer）和全连接层（Fully Connected Layer）等层次结构，从图像中提取特征，并进行分类或回归任务。

卷积层：通过卷积操作提取图像的局部特征，利用滤波器（卷积核）在图像上滑动，生成特征图（Feature Map）。卷积层的参数是可学习的，能够自动适应数据特点。

池化层：通常紧跟在卷积层后面，用于减少特征图的维度，从而降低模型计算量和防止过拟合。常见的池化操作包括最大池化（Max Pooling）和平均池化（Average Pooling）。

全连接层：将高维特征图展开为一维向量，并通过全连接的神经元进行分类或回归任务。全连接层类似于传统的神经网络，负责将提取的特征映射到输出空间。

CNN的优势在于能够自动提取图像的空间特征，减少人工特征工程的需求，并且在图像分类、目标检测等任务中表现出色。

### 3.3.2 CNN模型训练

在构建CNN模型时，本次设计了一个典型的卷积神经网络结构，包括以下几层：



图3-3 CNN模型构建代码

输入层：接受预处理后的图像数据，输入形状为（128, 128, 3），即128x128像素的彩色图像。

卷积层1：包含32个3x3的卷积核，激活函数采用ReLU。

池化层1：最大池化层，池化窗口大小为2x2，步幅为2。

卷积层2：包含64个3x3的卷积核，激活函数采用ReLU。

池化层2：最大池化层，池化窗口大小为2x2，步幅为2。

卷积层3：包含128个3x3的卷积核，激活函数采用ReLU。

池化层3：最大池化层，池化窗口大小为2x2，步幅为2。

全连接层：包含512个神经元，激活函数采用ReLU。

输出层：使用softmax激活函数，输出分类结果，类别数为垃圾分类的7个类别。

在模型训练过程中，采用以下参数配置：Adam优化器，学习率为0.001。交叉熵损失函数（categorical crossentropy）。批量大小32。训练轮数：20个epoch。

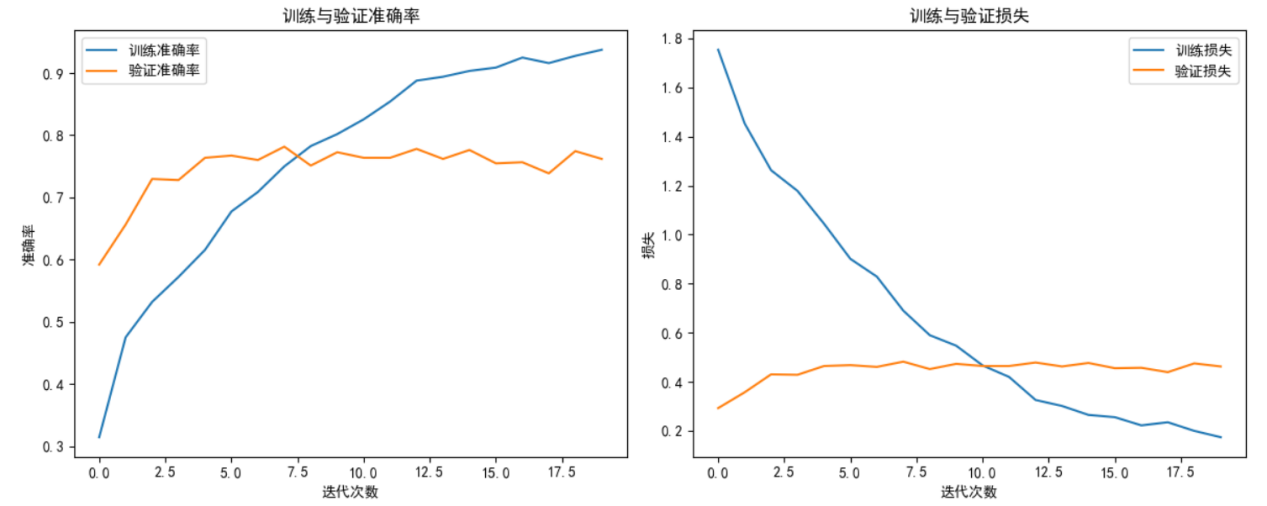


图3-4 CNN模型训练曲线图

训练损失和验证损失在前10个epoch内迅速下降，说明模型在初期能够快速学习到数据的特征。随着训练的进行，训练损失和验证损失趋于平稳，训练准确率和验证准确率逐步提高，并最终达到较高水平，表明模型的收敛效果良好。

## 3.4 模型结果对比评估

在垃圾图像分类任务中，我们构建了KNN、XGBoost和CNN三种模型，并对每个模型的准确率、精度、召回率和F1进行了评估。



图3-5 不同模型的评估结果

结果显示，CNN模型在各项指标上均表现优异，准确率为76.18%，精度为76.61%，召回率为76.18%，F1为75.67%。相比之下，XGBoost模型的表现次之，准确率为64.51%，精度为63.14%，召回率为64.51%，F1为61.20%。而KNN模型的表现最差，准确率为51.67%，精度为56.31%，召回率为51.67%，F1为51.16%。

通过对比分析，可以看出CNN模型在处理图像分类任务时具备更强的特征提取和分类能力，能够更准确地识别不同类别的垃圾图像。其较高的精度和召回率表明模型在分类过程中能有效捕捉到更多的正类样本，同时避免了较多的误分类情况。XGBoost模型虽然在传统机器学习模型中表现较好，但在处理高维图像数据时仍不如深度学习模型的表现。KNN模型由于其简单的原理和较低的计算复杂度，在面对复杂的图像分类任务时显得力不从心，无法提供足够的分类精度。因此，对于垃圾图像分类任务，CNN模型显然是最优选择。

# 4 结论

本次研究基于机器学习方法对垃圾图像进行分类，通过对KNN、XGBoost和CNN三种模型的构建和评估，全面比较了它们在垃圾分类任务中的表现。在数据收集和预处理阶段，我们将不同类别的垃圾图像统一调整为相同大小，并进行了图像归一化和数据增强，确保了数据集的多样性和均衡性。随后，通过合理的模型设计和参数调优，我们构建了KNN、XGBoost和CNN三种模型，并通过准确率、精度、召回率和F1等指标对模型进行了评估。

从实验结果来看，CNN模型在垃圾图像分类任务中表现最佳，其准确率、精度、召回率和F1指标均显著优于KNN和XGBoost模型。这表明CNN模型在处理高维图像数据和提取特征方面具有明显的优势，能够更有效地识别和分类不同类别的垃圾图像。相比之下，XGBoost模型虽然在传统机器学习模型中表现较好，但在面对复杂的图像数据时，仍不及深度学习模型的表现。KNN模型由于其简单的计算原理和较低的计算复杂度，在复杂的图像分类任务中显得力不从心。

综上所述，通过本次研究，我们验证了深度学习模型在垃圾图像分类任务中的优越性，特别是CNN模型在特征提取和分类方面的强大能力。未来的工作中，可以进一步优化CNN模型的结构和参数，探索更多的深度学习方法，如迁移学习和生成对抗网络等，以提高垃圾分类的精度和效率。同时，结合物联网技术，将垃圾图像分类系统应用于实际场景，如智能垃圾分类箱和垃圾处理系统，为智能城市建设和环境保护提供技术支持和解决方案。