

近红外光谱技术在常用包装材料分类方法的研究与应用

汪香君¹, 陈晨^{1*}, 李军¹, 熊露璐²

(1. 上海工程技术大学 材料科学与工程学院, 上海 201620;

2. 上海大觉包装制品有限公司, 上海 201706)

摘要: **目的** 探索近红外光谱技术在常见包装材料分类中的应用潜力。**方法** 首先利用近红外光谱仪对 PBS、HDPE、LDPE、LLDPE、PP、PS、POE、PPC、PBAT、PLA、PGA+PBAT 这 11 类常用包装材料进行光谱数据采集, 然后采用滑动平均 (MA)、标准正态变量变换 (SNV)、多元散射校正 (MSC) 对光谱进行预处理, 使用支持向量机 (SVM)、K-近邻 (KNN)、决策树和随机森林 4 种模式识别方法建立定性判别模型, 最后根据模型评价指标和混淆矩阵对模型的预测性能进行比较。**结果** SNV 结合 K-近邻的模型分类效果较优, 准确率达到 97.03%。**结论** 基于近红外光谱仪的塑料鉴别研究为塑料回收和再利用提供了一种便捷、快速、无损的检测方法, 具有较好的应用前景。

关键词: 近红外光谱技术; 包装材料; 分类

中图分类号: TB484.3

文献标志码: A

文章编号: 1001-3563(2024)17-0180-09

DOI: 10.19554/j.cnki.1001-3563.2024.17.022

Research and Application of Near-infrared Spectroscopy Technology in Classification Methods of Commonly Used Packaging Materials

WANG Xiangjun¹, CHEN Chen^{1*}, LI Jun¹, XIONG Lulu²

(1. School of Materials Science and Engineering, Shanghai University of Engineering Science, Shanghai 201620, China; 2. Shanghai Dajue Packaging Products Co., Ltd., Shanghai 201706, China)

ABSTRACT: The work aims to explore the potential of using near-infrared spectroscopy to classify common packaging materials. First, a NIR spectrometer was used to collect spectral data from 11 commonly used packaging materials, including PBS, HDPE, LDPE, LLDPE, PP, PS, POE, PPC, PBAT, PLA, and a blend of PGA+PBAT. The spectral data was then preprocessed using three techniques: Moving Average (MA), Standard Normal Variate (SNV) transformation, and Multiplicative Scatter Correction (MSC). Four pattern recognition methods, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree, and Random Forest were utilized to develop qualitative discrimination models. Finally, the prediction performance of the models was compared based on evaluation metrics and confusion matrices. The combination of SNV preprocessing with the KNN algorithm yielded the best classification performance, achieving an accuracy of 97.03%. It is concluded that NIR spectroscopy provides a convenient, fast, and non-destructive method for plastic identification, which is advantageous for plastic recycling and reuse. The results indicate good application prospects for this method in sustainable material management.

KEY WORDS: near infrared spectroscopy technology; packaging materials; classification

收稿日期: 2024-05-14

基金项目: 青科研专 2022-17; 上海工程技术大学材料学院高峰学科科研专项

*通信作者

塑料制品因其轻便、耐用、成本低廉等优点被广泛应用于制造业和日常生活。2008—2018 年, 全球塑料产量从 2.54 亿 t 增至 3.59 亿 t, 预计到 2050 年将增长 3 倍^[1]。然而, 塑料的大量使用和塑料废物的不当处理导致塑料废物在环境中积累^[2]。因此, 如何提高废旧塑料回收利用率成为亟待解决的问题, 而对塑料垃圾进行有序分类是提高回收利用率一大重要环节。

现阶段, 塑料的分类方法主要有机械分离法和自动分选法。机械分离方法包括密度分选^[3]、浮选法^[4]、静电分选^[5]等。这些方法具有成本低、效率高的优点, 但对塑料颗粒的尺寸大小、形状和表面条件都有一定的限制。此外, 机械分离方法在单次操作中仅能将混合物分成 2 个部分, 若要分离多组分废弃物, 需要通过多阶段分离或采用不同的方法组合进行处理。常用的自动分选技术包括近红外光谱法 (Near-Infrared, NIR)、中红外光谱法 (Mid-Infrared, MIR)、激光诱导击穿光谱法 (Laser-Induced Breakdown Spectroscopy, LIBS)、X 射线荧光法 (X-ray Fluorescence, XRF)、拉曼光谱法 (Raman Spectroscopy, RS) 等^[6-10]。这些技术可以同时实现 2 种以上材料的分离, 具有样品制备简单、操作便捷、信息获取快速等优点。美国材料试验学会 (American Society for Testing and Materials, ASTM) 将波长为 780~2 526 nm (12 820~3 959 cm^{-1}) 的电磁波定义为近红外光^[11]。该光谱区域与有机分子中含氢基团 (O—H、N—H 和 C—H) 的合频和倍频吸收区相同, 利用其吸收信息可对含氢基团物质进行定性定量检测。但由于吸收带高度重叠, 使得近红外光谱谱区解析十分困难。随着化学计量学技术以及计算机技术的发展, 近红外光谱数据分析更加快捷有效, 近红外光谱分析技术也逐渐进入了人们的视野之中。近年来, 近红外光谱技术也被广泛应用于塑料的自动分类。Kaihara 等^[12]采用近红外光谱技术, 运用决策树模型对 18 种塑料进行了分类。Huth-Fehre 等^[13]利用人工神经网络构建数据模型, 有效地区分了聚乙烯 (PE)、聚对苯二甲酸乙二醇酯 (PET)、聚丙烯 (PP)、聚苯乙烯 (PS) 和聚氯乙烯 (PVC) 5 种塑料。Rani 等^[14]通过主成分分析 (PCA) 和偏最小二乘判别分析 (PLS-DA), 对 5 种聚合物成功地进行了分类。张毅民等^[15]将近红外光谱技术与 Fisher 判别方法相结合, 建立了一种识别模型, 可以准确地鉴别 ABS、PET、PVC、PP、PS 和 PE 等废旧塑料, 其分类准确率达 100%。

本研究旨在基于近红外光谱技术对 PBS、HDPE、LDPE、LLDPE、PP、PS、POE、PPC、PBAT、PLA、PGA+PBAT 这 11 种常见塑料进行分类。通过截取波长范围、采用多种光谱预处理方法, 结合主成分分析及多种模式识别算法, 构建了用于包装材料材质快速鉴别的定性校正模型。通过对各个模型预测结果的评估与判断, 比选出最优模型, 为常用包装材料种类的

甄别提供参考, 为推动塑料回收再利用提供技术支持, 为更多应用场景的开发提供借鉴。

1 实验

1.1 材料与仪器

主要材料: 聚丁二酸丁二醇酯 (PBS)、高密度聚乙烯 (HDPE)、低密度聚乙烯 (LDPE)、线性低密度聚乙烯 (LLDPE)、聚丙烯 (PP)、聚苯乙烯 (PS)、聚烯烃弹性体 (POE)、聚碳酸亚丙酯 (PPC)、聚己二酸/对苯二甲酸丁二醇酯 (PBAT)、聚乳酸 (PLA)、聚乙交酯/聚己二酸/对苯二甲酸丁二醇酯 (PGA/PBAT), 采自上海大觉包装制品有限公司。

主要仪器: DZF-6050 真空烘箱, 广州市康恒仪器有限公司; SY-6210-C-20T 电热型压片机, 东莞市世研精密仪器有限公司; FLAME-NIR 近红外光谱仪, 蔚海光学仪器 (上海) 有限公司; HL2000-12 卤素光源、FIB-600-NIR 直通光纤、FIBH-2-UV 光纤准直镜、IS-20-5-T 透射积分球, 上海复享光学股份有限公司。

1.2 样品的制备

实验所需样品包含训练样品以及预测样品。训练样品由塑料颗粒熔融压膜制成。所有粒料均置于烘箱中以 60 °C 干燥 24 h。利用电热型压片机以及模具将若干粒料压成厚度为 2 mm 的薄膜, 设置一定的上下模腔硫化温度, 硫化时间 3 min, 加压 25 MPa, 每份粒料制备 1 个薄膜。将每个薄膜样品裁剪成 20 个相同尺寸 (20 mm×20 mm×2 mm) 的样品, 总样品数为 220 份。因一些着色添加剂在近红外光谱区域具有很强的吸收而难以识别^[16-17], 故本研究中的样品为透明或白色。

1.3 光谱采集

利用近红外光谱仪对塑料样品进行光谱采集。在正式采集光谱前需要对光源进行足够时间 (30 min) 的预热, 可以有效避免基线漂移对数据采集造成影响。如图 1 所示, NIR 光谱采样台的构成包括卤素光源、光谱仪、直通光纤、准直镜、透射积分球和计算机。光谱仪的扫描范围为 900~1 700 nm, 内含 128 个波段, 光谱分辨率约为 5.9 nm。设置积分时间为 800 ms、平均扫描次数为 15、滑动平均宽度为 5。光谱采集采用积分球透射模式, 样品与积分球球口之间的距离约为 3 mm。每个样品每隔 5 s 进行一次采集, 每次采集 10 条光谱, 共采集 20 次, 每个样品得到 200 条光谱, 总共得到 2 200 条光谱。所有光谱测量均使用空气作为参考背景。

1.4 数据处理软件

本研究的数据处理以及模型训练均在 Python 中完成。

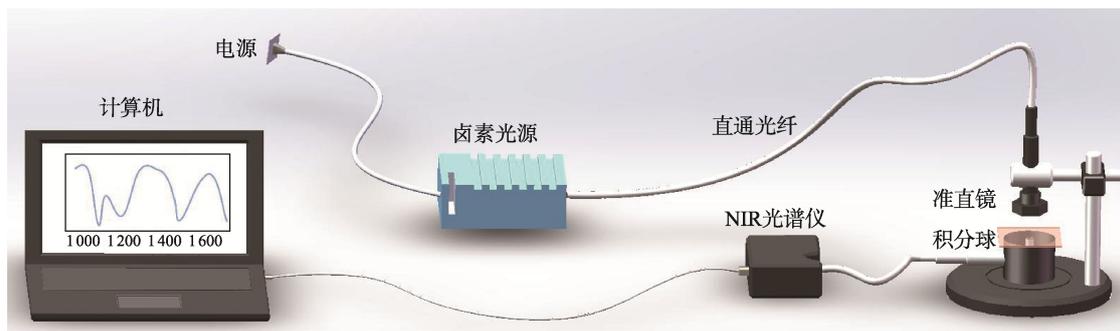


图1 近红外光谱采样台

Fig.1 Near-infrared spectroscopy acquisition platform

2 分类模型的构建

2.1 光谱预处理与样品集划分

塑料样品采用 K-S (Kennard-Stone) 方法^[18]进行样品集划分, 训练集与预测集比例为 3 : 1。每类样品有 150 条训练集光谱、50 条预测集光谱。光谱仪器的不稳定性、外界环境以及操作失误均会导致光谱数据异常, 从而影响模型的准确性和可靠性。本实验基于均值偏移百分比的方式剔除光谱数据中的异常值。计算每个光谱数据点与整体数据集的均值之间的偏移量, 并将偏移量与一定的阈值进行比较, 超过阈值的数据点被认为是异常的, 进而被剔除。

近红外光谱容易受到仪器精度、环境条件、样品性质等因素的综合影响。这些因素的变化可能导致光谱数据不稳定或出现偏差, 因此, 在进行光谱分析前, 通常需要对数据进行预处理, 以确保不同样本的光谱数据具有可比性^[19]。本研究采用了滑动平均 (Moving Average, MA)、标准正态变量变换 (Standard Normal

Variate Transformation, SNV)、多元散射校正 (Multiplicative Scatter Correction, MSC) 对光谱进行预处理, 比较结果见表 1。SVM 模型和 KNN 模型结合 SNV 预处理方法后均表现出较高的准确率, 这可能是因为 SNV 能够有效地消除光谱数据中的基线漂移和强度变化^[20], 使得模型更容易捕捉到样本之间的特征差异, 从而提高了模型分类准确率。决策树模型和随机森林模型在加入 SNV 预处理方法后的准确率不佳。经研究与分析, 其原因可能是对于这 2 种模型, SNV 不能有效地处理数据中的特征信息, 导致了信息的丢失, 进而降低了模型分类精度。使用 MSC 方法进行预处理后, 准确率有了明显提升, 经研究与分析, 可能是因为 MSC 可以有效地校正光谱数据中的多项式基线漂移, 使得决策树更加有效地捕捉到样本之间的特征差异。综合考虑 4 种模型在不同预处理方法下的结果, 本文采用 SNV-SVM、SNV-KNN、MSC-Decision Tree、MSC-Random Forest 的组合方式进行建模。

表 1 预测集样品分类模型的性能

Tab.1 Performance of classification model for prediction set samples

模型	预处理方法	准确率/%	精确率/%	召回率/%	F_1 分数($\times 100$)
SVM	SNV	99.27	99.36	99.27	99.27
	MSC	90.85	90.91	90.91	90.18
	MA	61.82	44.82	61.82	50.55
KNN	SNV	100	100	100	100
	MSC	91.15	91.27	91.18	91
	MA	52.06	31.91	52.09	38.64
Decision Tree	SNV	9.09	0.82	9.09	1.55
	MSC	77.39	79.18	77.36	77.64
	MA	47.21	33.55	47.18	36.09
Random Forest	SNV	9.09	0.82	9.09	1.55
	MSC	83.76	85.27	83.73	84
	MA	59.21	43.64	59.27	48.36

2.2 主成分分析

主成分分析 (Principal Component Analysis, PCA) 是一种无监督学习算法, 用于发现数据中的主要结构并减少数据的复杂性^[21]。它通过线性变换将原始数据映射到新的坐标系中, 使得数据在投影后具有最大的方差, 通常用于降维、数据可视化和去噪等方面。本研究采用主成分分析法评估光谱数据的聚类趋势。将光谱压缩成一组线性不相关的变量, 称之为主成分 (PCs)^[22]。第 1 个 PC 解释了红外光谱的最大变异性, 随后的 PC 解释了不断减少的变异性^[23]。

2.3 模型分类方法

支持向量机 (Support Vector Machine, SVM) 是一种建立在核方法和统计学习理论的监督学习模型, 可应用于分类、回归及检测数据中的异常值问题^[24]。该模型具有较强的学习能力, 能够在有限且复杂的样本信息中找到最佳的数学模型。其基本原理是首先将输入的向量通过核函数映射到高维特征空间, 然后在特征空间上寻找一个最优的超平面, 以最大化不同类别样本之间的间隔, 并将它们有效地分开^[25]。与其他分类方法相比, SVM 分类器在优化过程中使用损失函数和惩罚项来平衡间隔最大化和误分类点的惩罚, 以找到最优的超平面, 从而实现对本样本的分类。本实验采用网格搜索 (Grid Search) 的方式对 SVM 模型进行超参数调优, 首先通过尝试一系列指定的超参数组合, 其次利用交叉验证对每个组合进行评估, 最终确定性能最优的超参数组合。

随机森林 (Random Forest, RF) 是一项基于集成学习的机器学习技术, 它的核心思想是通过构建多个决策树来实现分类和回归任务。模型引入了随机性的元素, 即随机选择样本和特征, 从而增加了模型的多样性, 减少了过拟合的风险^[26]。随机森林模型由决策树和 2 种随机选择策略组成。随机选择样本即随机有放回地从原始数据集中抽取若干样本, 构成一个新的样本集合, 称为自助样本集。使用自助样本集来建立决策树, 随机选择特征即在决策树的构建过程中, 在树的每个节点处, 从所有特征中随机选择一个固定数量的特征子集。最后集成多个决策树的预测结果, 通过统计学的方法进行输出。

K-近邻 (K-Nearest Neighbor, KNN)^[27-28]是一种基于实例学习的机器学习算法, 适用于监督学习环境中的分类问题。其主要思想是利用数据集中与新样本距离最接近的 K 个邻居的标签来预测新样本的标签。常见的距离度量方法包括欧氏距离、曼哈顿距离等。KNN 算法简单直观, 无需训练过程, 且不需要假设数据的分布形式, 因此在一些简单的分类和回归问题上表现良好。但在大规模数据和高维数据上的计算复杂度较高, 且对不平衡数据集容易产生分类偏差。

决策树 (Decision Tree, DT)^[29-30]是一种基于树形结构的分类与回归模型, 通过一系列的特征选择和划分节点, 将数据集划分成具有相似特征子集, 直到达到某个终止条件为止。其中根节点代表最初的判定依据, 中间节点表示进一步的判定步骤, 而叶子节点则代表最终的决策结果。它具有直观性和易解释性, 适用于处理非线性关系和高维数据, 但容易产生过拟合等问题。决策树常用的划分准则有信息增益、信息增益率和基尼指数。不同的准则在不同的数据集和任务上表现不同, 因此在实际应用中需要根据具体情况选择合适的划分准则。

2.4 模型评价指标

交叉验证 (Cross Validation, CV) 是估计预测精度的一种简单且常用的方法, 用于评估模型在新数据上的泛化性能^[31]。它将数据集划分为多个子集, 通过反复训练和测试模型, 从而减少了因数据集划分不合理而引入的偏差, 提高了模型评估的可靠性。为避免随机数据分割的影响, 本文采取 10 倍交叉验证方法, 获得平均值。

模型评价采用混淆矩阵以及多分类问题的性能评价指标。混淆矩阵 (Confusion Matrix) 又称误差矩阵, 一种在机器学习和统计学中用于评估分类模型性能的形式。它将模型的预测结果与实际标签进行比较, 以便更全面地了解模型在不同类别上的性能表现^[32]。评价指标包括准确率 ACC (Accuracy)、精密率 P (Precision)、召回率 R (Recall) 和 F_1 分数, 公式见 (1)~(4)。以上 4 个数值越高, 则说明模型的识别能力越好。

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (1)$$

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (2)$$

$$F_1 = \frac{2P \times R}{P + R} = \frac{2T_p}{2T_p + F_p + F_n} \times 100\% \quad (3)$$

$$A_c = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \times 100\% \quad (4)$$

式中: T_p (True Positive) 为模型正确地将正类样本预测为正类的数量; F_p (False Positive) 为模型错误地将负类样本预测为正类的数量; T_n (True Negative) 为模型正确地将负类样本预测为负类的数量; F_n (False Negative) 为模型错误地将正类样本预测为负类的数量; A_c 为模型正确预测的样本数占总样本数的比例。

3 结果与讨论

3.1 塑料样品的近红外光谱

11 类原始塑料的近红外透射光谱如图 2 所示。被检塑料在 1 200 nm 和 1 400 nm 附近均有吸收峰,

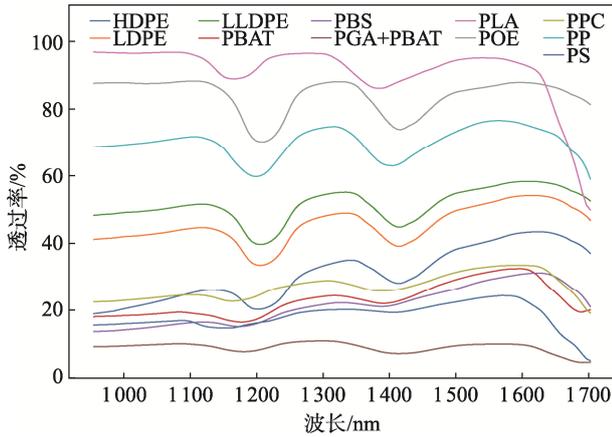


图2 塑料的平均近红外光谱
Fig.2 Average near-infrared spectrum of plastics

这分别对应的是C—H键的二级倍频峰和C—H键的组频峰。

经研究与分析,由于谱线上的明显的区别,PLA和其他塑料更易区分。而LLDPE和LDPE的近红外光谱非常相似,在900~1700 nm的范围内吸收峰的个数一致,位置也十分接近,难以从图谱直接分辨,这一点也在后续的分类中得到了证实。因此,需通过化学计量学方法对光谱进行进一步的判别。

为了从原始数据中提取出最具信息量的特征,降低数据的维度,减少计算成本以及提高模型的训练效率和泛化能力,只利用部分波段光谱进行建模。

从光谱图中可以看出,大多数材料在900~1100 nm区域吸收较弱,并且在接近最大波长时仪器噪声较强。根据一阶导数光谱,选取11个光谱在1100 nm附近的峰起点和1600~1700 nm的极值点,将2组波长平均后,分别选择1096.39 nm和1618.73 nm作为起始和结束波长。

3.2 塑料样品近红外光谱主成分分析

为了更好地评估模型的泛化能力,从市场购买塑料膜袋、容器以及板材作为验证样品。将前2个主成分设置为变量,其二维PCA评分图如图4所示。通过这种可视化方式,可以清晰地展示原始样本和验证样本的光谱数据结构,进而比较它们之间的差异。在原始样品的散点图中,PLA、PP、POE和PBS的数据结构聚为不同的组,LDPE和LLDPE以及PS、PPC和PBAT的聚类虽非常接近,然而仍然可以看到差异。在验证样品的图中,数据点更加分散,说明样品的多样性导致了同一类型的光谱之间的差异更大。在这11种类型中,PPC和HDPE、PBS和PBAT的簇相对接近,LDPE和LLDPE的数据由于光谱的相似性,表现为几乎在同一区域分布。尽管PCA在数据探索方面表现出色,但由于其属于无监督学习模式,不能直接用于解决多种塑料样品的分类问题。为了更准确地区分不同类型的塑料样品,通过采用有监督学习的模式识别方法来构建有效的分类模型。

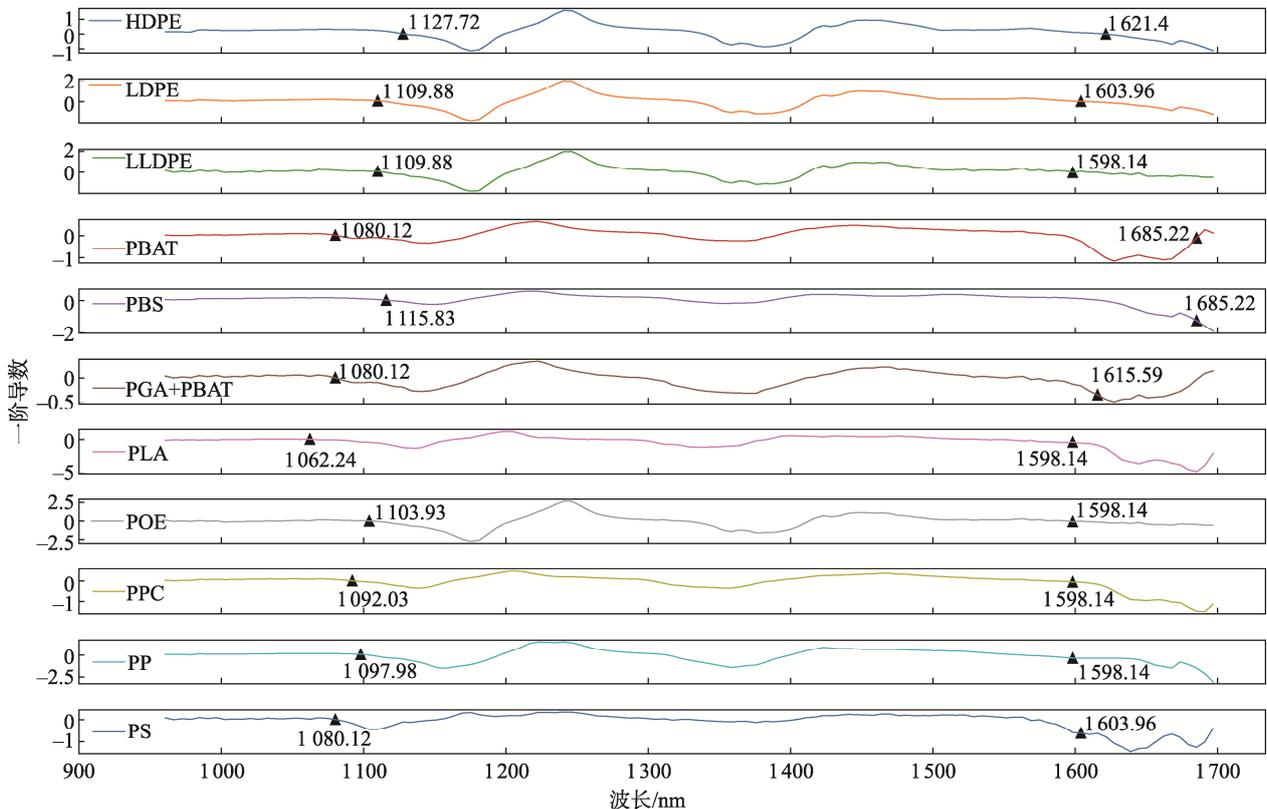


图3 11种塑料的一阶导数光谱
Fig.3 First derivative spectra of 11 plastics

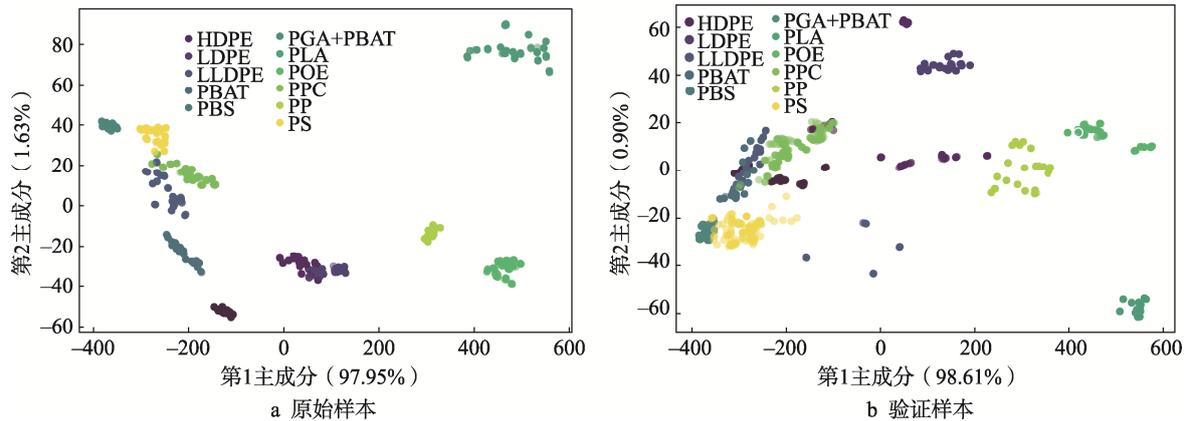


图 4 光谱数据 PCA 评分图
Fig.4 PCA score plot of spectral data

3.3 分类模型的比较与讨论

为了研究这些方法对塑料分类的可行性, 将原始 PBS、HDPE、LDPE、LLDPE、PP、PS、POE、PPC、PBAT、PLA 和 PGA+PBAT 的光谱按照 K-S 方法划分

为训练集 (75%) 和测试集 (25%)。对训练数据进行交叉验证, 选择最佳预处理方法, 在剩余测试集上进行预测。最终的预测结果如图 5 所示, 与 MSC-Decision Tree 和 MSC-Random Forest 模型相比, SNV-SVM 和 SNV-KNN 模型的分类效果较好,

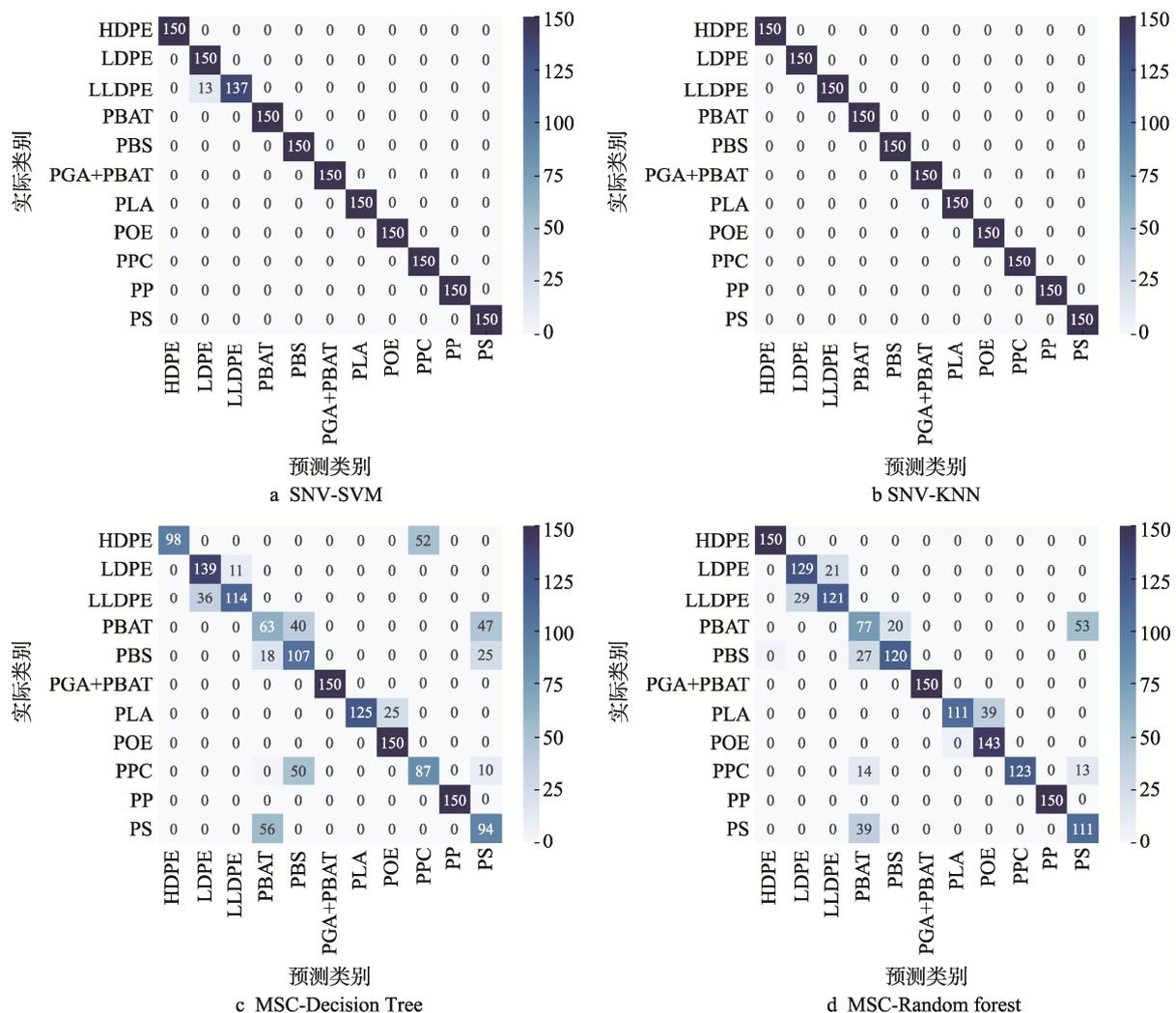


图 5 原始塑料预测集混淆矩阵
Fig.5 Confusion matrix for original plastic prediction set

SNV-KNN 的准确率达到 100%，SNV-SVM 的准确率也接近 100%。一些光谱非常相似的塑料，如 LDPE 和 LLDPE，也能被很好地识别出来。这说明 SNV-KNN 模型能够准确区分这些原生塑料。

为了探索模型利用从原生塑料中学习的知识对废塑料进行分类的能力，将测试集替换为验证样本（共 2 200 个测量值），训练集仍由原生塑料组成。4 种方法的混淆矩阵如图 6 所示，PLA 和 PGA+PBAT 可显著地与其他塑料区分开来，与前述研究呼应。对

于 SVM 模型，149 个 LDPE 样品被鉴定为 LLDPE。而在决策树模型中，14 个 LDPE 样品被鉴定为 LLDPE，剩下的被预测为 PPC 和 PP。因此，决策树模型在塑料分类领域的应用非最优方案。

将 4 种分类方法的预测结果用准确率、召回率、F1 分数评价指标来进行比较，结果见表 2。SNV-KNN 的准确率最高，达到 97.03%，精确率和召回率分别为 97.36%、97%， F_1 评分为 97.09。因此，对于以上塑料样品，选择 SNV-KNN 作为最优模型，其次是支

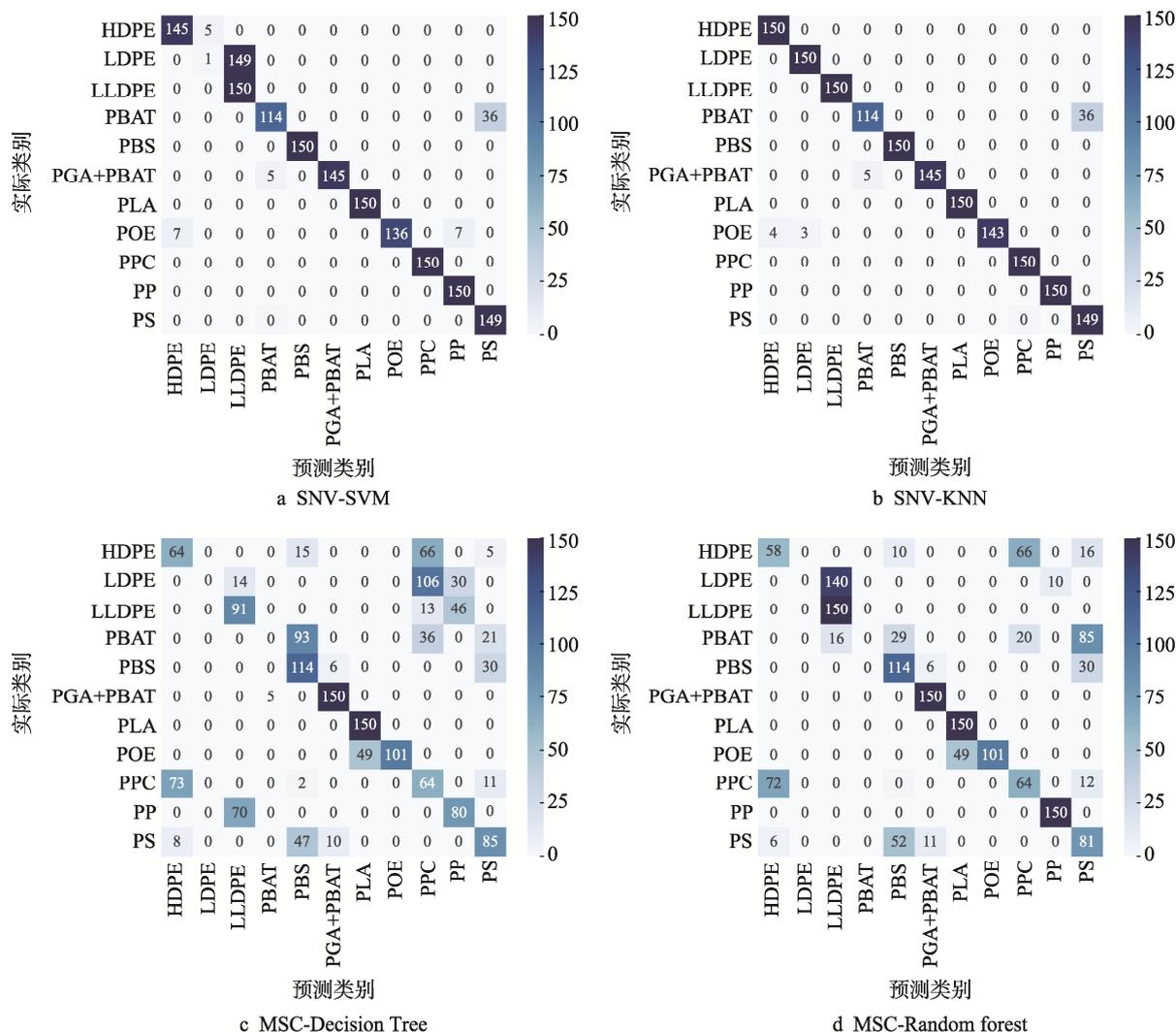


图 6 外部验证集混淆矩阵
Fig.6 Confusion matrix for external validation set

表 2 验证集样品分类模型的性能
Tab.2 Performance of classification model on validation set samples

模型	预处理方法	准确率/%	精确率/%	召回率/%	F_1 分数 (×100)
SVM	SNV	87.27	84.91	87.36	84.36
KNN	SNV	97.03	97.36	97	97.09
Decision Tree	MSC	54.48	48.36	53.91	50.09
Random Forest	MSC	61.70	53.18	61.73	55.91

持向量机模型。二者都能达到较好的精度,并且在高维空间或者数据集结构具有复杂结构时,SVM可能更佳,因为它可以更好地处理非线性问题。决策树和随机森林模型的准确率不佳,因其为基于树结构的模型,在处理复杂数据集时可能会受到过拟合的影响。特别是在特征空间较大或者数据集较小的情况下,决策树和随机森林可能会倾向于过于细致地划分数据,导致对训练数据过拟合,从而影响了在未见数据上的泛化能力。

4 结论

本研究基于近红外光谱仪,通过机器学习方法建立了 PBS、HDPE、LDPE、LLDPE、PP、PS、POE、PPC、PBAT、PLA、PGA+PBAT 这 11 类常用塑料的鉴别模型,探索了常用塑料的鉴别方法,并取得了一定的效果。结果表明,SNV-KNN 模型的总体准确率、精确率和召回率最高,预测集准确率为 100%,外部验证集准确率也达到 97.03%,且该模型受训练数据与测试数据差异的影响较小,可以实现 11 种塑料的一次性识别。基于上述研究的结果得出,近红外光谱技术在常用塑料的鉴别方面表现出了潜在的应用前景。选择适当的预处理和多元变量分析方法,不仅可用于识别标准塑料样品,也可以用于常用塑料制品。

然而,使用近红外光谱进行塑料分选鉴别仍然有一些局限性。该技术通常只适用于浅色塑料,因为深色塑料吸收光线较多,难以准确识别,且需要大量样本数据进行训练以获得准确的模型。因此,后续需要进一步优化算法以及采集更多样本数据,以提高鉴别的准确性和可靠性。综上所述,基于近红外光谱仪的塑料鉴别研究为塑料回收和再利用提供了一种便捷、快速、无损的检测方法,不仅对当下塑料污染治理起到有效支撑,还具有重要的应用前景和发展价值。

参考文献:

- [1] OTHMAN A R, ABU HASAN H, MUHAMAD M H, et al. Microbial Degradation Of microplastics by Enzymatic Processes: A Review[J]. Environmental Chemistry Letters, 2021, 19(4): 3057-3073.
- [2] KURNIAWAN T A, HAIDER A, AHMAD H M, et al. Source, Occurrence, Distribution, Fate, and Implications of Microplastic Pollutants in Freshwater on Environment: A Critical Review and Way Forward[J]. Chemosphere, 2023, 325: 138367.
- [3] LU T F, CHEN W T. Material Recycling of Acrylonitrile Butadiene Styrene (ABS) from Toy Waste Using Density Separation and Safer Solvents[J]. Resources, Conservation and Recycling, 2023, 197: 107090.
- [4] KÖKKİLİÇ O, MOHAMMADI-JAM S, CHU P B, et al. Separation of Plastic Wastes Using Froth Flotation-an Overview[J]. Advances in Colloid and Interface Science, 2022, 308: 102769.
- [5] ACHOURI I E, BOUKHOULDA M F, MEDLES K, et al. Electrostatic Separation of Tribocharged Granular Mixtures of Two or More Plastics Originating from WEEE[J]. IEEE Transactions on Industry Applications, 2022, 58(6): 7701-7708.
- [6] XIA J J, HUANG Y, LI Q Q, et al. Convolutional Neural Network with Near-Infrared Spectroscopy for Plastic Discrimination[J]. Environmental Chemistry Letters, 2021, 19(5): 3547-3555.
- [7] LONG F, JIANG S L, ADEKUNLE A G, et al. Online Characterization of Mixed Plastic Waste Using Machine Learning and Mid-Infrared Spectroscopy[J]. ACS Sustainable Chemistry & Engineering, 2022, 10(48): 16064-16069.
- [8] WANG Q Q, HUANG Z W, LIU K, et al. Classification of Plastics with Laser-Induced Breakdown Spectroscopy Based on Principal Component Analysis and Artificial Neural Network Model[J]. Spectroscopy & Spectral Analysis, 2012, 32(12): 3179-3182.
- [9] ABUBAKER S A, TAHA A H. Identification and Characterization of Different Types of Plastics Wastes Using X-ray Diffraction and X-ray Fluorescence Techniques[J]. ARO-The Scientific Journal of Koya University, 2021, 9(2): 22-25.
- [10] YANG Y, ZHANG W, WANG Z, et al. Differentiation of Plastics by Combining Raman Spectroscopy and Machine Learning[J]. Journal of Applied Spectroscopy, 2022, 89(4): 790-798.
- [11] REICH G. Near-Infrared Spectroscopy and Imaging: Basic Principles and Pharmaceutical Applications[J]. Advanced Drug Delivery Reviews, 2005, 57(8): 1109-1143.
- [12] KAIHARA M, SATOH M, SATOH M. Systematization Method for Distinguishing Plastic Groups by Using NIR Spectroscopy[J]. Analytical Sciences, 2007, 23(7): 921-924.
- [13] HUTH-FEHRE T, FELDHOFF R, KANTIMM T, et al. NIR - Remote Sensing and Artificial Neural Networks for Rapid Identification of Post Consumer Plastics[J]. Journal of Molecular Structure, 1995, 348: 143-146.
- [14] RANI M, MARCHESI C, FEDERICI S, et al. Miniaturized Near-Infrared (MicroNIR) Spectrometer in Plastic Waste Sorting[J]. Materials, 2019, 12(17): 2740.
- [15] 张毅民, 白家瑞, 刘红莎, 等. 基于近红外的 Fisher 判别法鉴别废塑料[J]. 工程塑料应用, 2014,

- 42(5): 75-79.
- ZHANG Y M, BAI J R, LIU H S, et al. Recognition of Plastics Wastes with Fisher Discriminant Classification Based on near Infrared Spectrum[J]. *Engineering Plastics Application*, 2014, 42(5): 75-79.
- [16] MARTÍNEZ S S, PANIZA J M L, RAMÍREZ M C, et al. A Sensor Fusion-based Classification System for Thermoplastic Recycling[C]// 18th International Conference on Automation and Computing (ICAC). IEEE, 2012: 1-6.
- [17] ROZENSTEIN O, PUCKRIN E, ADAMOWSKI J. Development of a New Approach Based on Midwave Infrared Spectroscopy for Post-Consumer Black Plastic Waste Sorting in the Recycling Industry[J]. *Waste Management*, 2017, 68: 38-44.
- [18] KENNARD R W, STONE L A. Computer Aided Design of Experiments[J]. *Technometrics*, 2012, 11(1): 137-148.
- [19] ROBERT G, GOSELIN R. Evaluating the Impact of NIR Pre-Processing Methods *via* Multiblock Partial Least-Squares[J]. *Analytica Chimica Acta*, 2022, 1189: 339255.
- [20] KHODABAKHSHIAN R, SEYEDALIBEYK LAVASANI H, WELLER P. A Methodological Approach to Preprocessing FTIR Spectra of Adulterated Sesame Oil[J]. *Food Chemistry*, 2023, 419: 136055.
- [21] DONG W, WOŹNIAK M, WU J, et al. Denoising Aggregation of Graph Neural Networks by Using Principal Component Analysis[J]. *IEEE Transactions on Industrial Informatics*, 2022, 19(3): 2385-2394.
- [22] WANG W, WANG C C, SHEN J N, et al. Aggregate Fingerprints Identification Based on Its Compositions and Machine Learning Algorithm[J]. *Arabian Journal of Chemistry*, 2023, 16(7): 104810.
- [23] JETYBAYEVA A, BORODINOV N, IEVLEV A V, et al. A Review on Recent Machine Learning Applications for Imaging Mass Spectrometry Studies[J]. *Journal of Applied Physics*, 2023, 133(2): 020702.
- [24] CORTES C, VAPNIK V. Support-Vector Networks[J]. *Machine Learning*, 1995, 20(3): 273-297.
- [25] ALFALEH A, BEN KHEDHER N, ELDIN S M, et al. Predicting Thermal Conductivity and Dynamic Viscosity of Nanofluid by Employment of Support Vector Machines: A Review[J]. *Energy Reports*, 2023, 10: 1259-1267.
- [26] IMMITZER M, VUOLO F, ATZBERGER C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe[J]. *Remote Sensing*, 2016, 8(3): 166.
- [27] HATTORI K, TAKAHASHI M. A New Nearest-Neighbor Rule in the Pattern Classification Problem[J]. *Pattern Recognition*, 1999, 32(3): 425-432.
- [28] 宋毅飞, 周剑秋. KNN算法与其改进算法的性能比较[J]. *机电产品开发与创新*, 2017, 30(2): 60-63.
- SONG Y F, ZHOU J Q. Performance Comparison of KNN Algorithm and Its Improved Algorithm[J]. *Development & Innovation of Machinery & Electrical Products*, 2017, 30(2): 60-63.
- [29] IMANI M, ARABNIA H R. Hyperparameter Optimization and Combined Data Sampling Techniques in Machine Learning for Customer Churn Prediction: A Comparative Analysis[J]. *Technologies*, 2023, 11(6): 167.
- [30] 赵涓涓, 马瑞良, 张小龙. 基于决策树和改进 SVM 混合模型的语音情感识别[J]. *北京理工大学学报*, 2017, 37(4): 386-390.
- ZHAO J J, MA R L, ZHANG X L. Speech Emotion Recognition Based on Decision Tree and Improved SVM Mixed Model[J]. *Transactions of Beijing Institute of Technology*, 2017, 37(4): 386-390.
- [31] 杨柳, 王钰. 泛化误差的各种交叉验证估计方法综述[J]. *计算机应用研究*, 2015, 32(5): 1287-1290.
- YANG L, WANG Y. Survey for Various Cross-Validation Estimators of Generalization Error[J]. *Application Research of Computers*, 2015, 32(5): 1287-1290.
- [32] RIEHL K, NEUNTEUFEL M, HEMBERG M. Hierarchical Confusion Matrix for Classification Performance Evaluation[J]. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 2023, 72(5): 1394-1412.